National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

THE UNIVERSITY OF ALBERTA

Scaling Models for the Measurement of Infant Motor Development

by

Lynn Ellen Pinnell

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

Edmonton, Alberta

FALL 1991

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-70006-8

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Lynn Ellen Pinnell

TITLE OF THESIS: Scaling Models for the Measurement of Infant Motor Development

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1991

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Lynn Pinnell

15103 - 56 Avenue

Edmonton, Alberta

T6H 5B3

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled Scaling Models for the Measurement of Infant Motor Development submitted by Lynn Ellen Pinnell in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. T. Maguire

Dr. S. Hunka

Dr. M. Piper

Dr. L. Stewin

Dr. M. Bouffard

Dr. L. Fetters

Date October 8, 1991

To my children, Eric, Ian and Mark,

and to my best friend, Paul

Abstract

This research was part of a three-year process of development and validation of the Alberta Infant Motor Scale, an instrument for assessing gross motor maturation in infants from birth to the age of independent walking. The purposes of the research were to investigate and recommend a model for the scaling of items on this instrument and to recommend a scoring system for the final scale. The sample consisted of 506 normal infants and was age-stratified through the first 18 months of life. Tests of dimensionality of the 58 item set provided strong evidence for a single dimension underlying the data. Several models were tested for the scaling of items along the single dimension, and all produced consistent results. Multidimensional scaling was recommend as the most useful approach for this purpose. Two scoring systems were explored with these data and some tentative recommendations made, pending further data on abnormal and high-risk infants.

Acknowledgment

The author wishes to express sincere appreciation to the members of the thesis committee for their thoughtful contributions to this work, and particularly to Dr. Tom Maguire, the thesis supervisor, for his consistent guidance and encouragement over the past four years. Dr. Martha Piper is acknowledged for her support of the thesis research and for her outstanding leadership as the principal investigator of the Alberta Infant Motor Scale Project.

The author also expresses thanks to the following individuals: Annette Kujda, for coordinating the data collection and data entry; the physical therapists who collected the data, particularly Johanna Darrah who shared her content expertise at several stages of the study; the faculty and graduate students in the Centre for Research in Applied Measurement and Evaluation for their interest in the research, and especially to Paul Beaulne for his ideas and assistance regarding the data analysis.

Gratitude is expressed to the parents and infants who participated in the study and to the Edmonton Board of Health, which provided access to the infants.

The Alberta Infant Motor Scale project, of which this study was a part, was funded by the National Health Research and Development Programs, Health and Welfare Canada, in a grant to Drs. Martha Piper, Tom Maguire, Paul Byrne and Tom

Paton. NHRDP is further acknowledged for support of the author as a Ph.D. Fellow during the time that this research was completed. The author also thanks the University of Alberta for its assistance in the form of the Dissertation Fellowship and other awards.

Table of Contents

Chapter 1 1

Introduction 1

Background to the Problem 1

Prior Work Completed by the Research Team . . 3

Rationale for the Scaling Study 6

Chapter 2 9

Review of Selected Literature 9

Principles and Models of Early Motor Development 9

Review of Scaling Models 19

The Issue of Dimensionality 22

Multidimensional Scaling 23

Factor Analysis 24

Item Response Models 25

Guttman Scaling 26

The Structural Validity Context 27

Chapter 3 30

Methods 30

Sample 30

Design 31

Data Collection Procedures 32

Scoring of Items 33

Data Analysis 35

Tests of Dimensionality 36
Methods for Scaling Items 37
Item Discriminations 39
Determination of Scoring System 40

Chapter 4 41
Results and Discussion 41
Descriptive Statistics - Infants 41
Descriptive Statistics - Items 41
Dimensionality of the Data Set 42
Multidimensional Scaling Results 42
Factor Analysis Results 45
Guttman Scaling Results 46
Scaling of the Items 47
Item Discrimination Estimates 50
Scoring System 51

Chapter 5 55
Conclusions and Recommendations 55

References 58

Appendices 76

List of Tables

Table	Description	Page
1.	Demographic Characteristics of the Sample . . .	63
2.	Age Grouping of Infants	64
3.	Score Frequencies for Items	65-66
4.	Multidimensional Scaling of 58 Items	67
5.	Item Scale Values from MDS	68
6.	Multidimensional Scaling by Group.	69
7.	Factor Analysis (NOHARM) Results	70
8.	Ordering of Items by Various Models	71-72
9.	Age Placement of Items	73
10.	Correlations Between Item Placement Estimates. .	74
11.	Discrimination Estimates from Various Models . .	75

Chapter 1

Introduction

The purposes of this study were to investigate and recommend a model to be used for the scaling of the Alberta Infant Motor Scale, a new assessment instrument to measure infant motor development, and to recommend a scoring system for this instrument.

Background to the Problem

In July, 1988, a group of researchers at the University of Alberta was awarded a grant from the National Health Research and Development Programs to construct and validate a new instrument for the assessment of infant motor development from birth (forty weeks gestation) through the age of independent walking. This instrument, the Alberta Infant Motor Scale (AIMS) is intended for use by physical therapists whose practice involves the follow-up and assessment of high-risk infants. It should provide a more sensitive measure than is currently available for the early identification of motor problems and for determining the efficacy of treatment of such problems.

The few standardized measures of infant motor development that do exist have been designed to measure motor development only in terms of the attainment of major motor milestones (Bayley, 1969; Folio and Dubose, 1974; Griffiths, 1954; Wolanski, et al, 1973). Physical therapists who work with these instruments view them as very

gross measures of motor performance and believe that the quantitative assessment of motor milestones has limited value in detecting early signs of dysfunction or in detecting qualitative changes in motor performance over time.

Physical therapists working in neonatal follow-up clinics and treatment programs for at-risk infants are hampered by the absence of a reliable and valid measure of early motor development. As well, efforts to assess the efficacy of early intervention programs for at-risk infants are seriously deterred by the lack of appropriately scaled, standardized measures of motor development. Without a sensitive outcome measure, costly treatment programs aimed at enhancing the motor development of at-risk infants will remain unaccountable.

The overall objectives of the AIMS are: 1) to identify infants whose motor performance is delayed or aberrant relative to a normative group, 2) to provide parents and clinicians with information about the motor activities the infant has mastered, those currently developing, and those not yet in the infant's repertoire, 3) to measure motor performance over time or before and after intervention, 4) to measure changes in motor performance that are quite small and thus not likely to be detected using more traditional instruments, and 5) to be an appropriate

research tool for assessing the efficacy of rehabilitation programs for infants with motor disorders.

Prior Work Completed by the Research Team

The construction and validation phases of the project have taken place over three years, and the steps are briefly summarized below.

1. A total of 84 items was initially generated, based upon published descriptive narratives of early motor performance. Four sets of items were written, corresponding to the four positions in which infants are to be assessed: prone, supine, sitting and standing. Each item consists of a drawing of an infant in a particular position, accompanied by a detailed description of the weight-bearing, posture and antigravity movements observed in that position. All items are scored on a pass/fail basis; that is, an infant must correctly demonstrate all three components of the particular behaviour in order to pass the item. See Appendix A for some sample items.

2. An initial review of items was carried out by several pediatric physical therapists who work with infants at the Glenrose Provincial General Hospital (Edmonton) and the Alberta Children's Hospital (Calgary).

3. Input from 118 pediatric physical therapists across Canada was received in response to a mail inquiry sent to 291 therapists. The respondents rated each item as to its

importance to motor development, the likelihood that an infant would demonstrate that behaviour during an assessment, and the observability of the behaviour, if demonstrated. They were also asked to sort the items according to their typical order of emergence, and to give an age range within which each behaviour would be expected to emerge in normal infants. The analysis of responses resulted in the elimination of 17 items, and the revision of certain others. Initial placement of remaining items along the continuum for motor development was accomplished using the therapists' averaged estimates of the ages of emergence in normal infants. In addition, the data from the item sorting task were subjected to a multidimensional scaling procedure to assess whether other dimensions besides the developmental sequencing one were necessary to account for the therapists' responses. Within each of the four subsets of items, a unidimensional model provided an adequate fit to the data, and this single dimension appeared to be developmental sequence.

4. A two-day work session was held with six international experts in infant motor development. The session was comprised of four stages. First, the experts were each given a copy of the item sets and were asked to review them for clarity, significance, order and inclusiveness. Second, they were asked to review in detail

certain item sets to determine the accurate sequence of items within the set, to remove inappropriate items, and to add any omitted items to the set. The third stage involved combining the four sets of items on a maturational continuum. Finally, a group session was held to discuss issues related to administration and scoring.

5. The instrument was revised and administration guidelines were developed in light of the input received from the work session. A number of items were deleted and a few added, resulting in a total of 59 items. A score sheet was developed for use in the feasibility testing.

6. A feasibility test was carried out using a sample of 97 normal infants, age-stratified through the first 18 months of life and recruited through the Edmonton Board of Health well-baby clinics. The data gathered led to recommendations for revisions of the instrument prior to reliability and validity testing. Specifically, seven items were deleted and six new items added resulting in a total of 58 items (21 prone, 9 supine, 12 sitting, and 16 standing). From the feasibility testing, it was determined that the assessment takes 15 to 20 minutes to complete, is easily scored through observation with little or no handling, is easily administered in a well-baby clinic, and requires minimal space and little special equipment.

Certain scaling models were tested on the feasibility data, including multidimensional scaling, Guttman scaling, and item response models. Although these analyses involved a very small number of infants per age category, the items appeared to be measuring a single dimension and the ordering of items on the developmental continuum was very close to what had been anticipated.

Rationale for the Scaling Study

An important part of the validation of the new instrument involves a search for the structure in the data obtained from a sample of normal infants, and an assessment of the extent to which this structure corresponds to accepted theoretical views of early motor development. Issues such as the number of dimensions or constructs underlying the data, the sequencing of items, and the scaled distances between items must be addressed if the new instrument and its scoring system are to provide for an accurate representation of motor ability as it really exists in normal infants.

In this regard, a model of scaling has a bridging function between the data, on the one hand, and the substantive theory, on the other (Van der Ven, 1980). Scaling may be defined as the attempt to find a set of coherent rules whereby non-physical objects (or properties thereof) can be represented by their position on a numerical

scale (Davies and Coxon, 1982). Van der Ven (1980) defines a scaling model as a numerical relational system, and measurement as the representation of an empirical relational system in or by a numerical relational system.

While the terms measurement and scaling are sometimes used interchangeably, we also often think of measurement as the process of obtaining a score for an individual on some instrument. In Torgerson's view, the logic of measurement concerns "the process by which the yardstick is developed, and not to its use once it has been established..." (1958, p. 14). This definition is consistent with conceptions of scaling.

Measurement is a process involving both theoretical and empirical considerations. From an empirical standpoint, the focus is on the observable behaviours; theoretically, our interest lies in the unobservable trait or construct represented by those behaviours. In this study, the unobservable construct (i.e., latent trait) of interest is gross motor maturity. The observable indicators of this trait are presumed to be the scores on the Alberta Infant Motor Scale. The task of scaling is to assign numerical values to the items on the instrument so that the representation of the underlying trait by the scores will be as meaningful as possible.

Based upon all of the earlier work in item generation and content validation, the research team (of which this writer was a member) was reasonably confident that the Alberta Infant Motor Scale measured a single construct, gross motor maturity. However, it was possible that certain subsets of items, particularly those appropriate for older infants, might also measure some other dimension such as cooperativeness. Also, since we had developed items within four postural positions (prone, supine, sitting, and standing), it would not be too surprising if items from these four positions separated out as measuring somewhat different dimensions of gross motor maturity. It is well known among physical therapists who work with abnormal and high-risk infants, that these infants will often perform better in certain positions than they do in others. Still, one would expect that normal infants would have age-appropriate performance across positions.

In summary, then, the primary purpose of this study was to investigate and recommend a model for the scaling of the fifty-eight items of the Alberta Infant Motor scale. The expectation was that, when applied to normal infants, these items would be scalable along a single dimension, and that this dimension would be gross motor maturity.

Chapter 2

Review of Selected Literature

The following literature review is organized into two sections. The first is concerned with theoretical views of early motor development, while the second is a review of scaling models.

Principles and Models of Early Motor Development

Most of the published literature on early motor development is in the form of descriptive accounts of the sequences of motor behaviours, often called developmental schedules. These accounts are usually based upon the early observational studies by Shirley (1931), Gesell (1940), and McGraw (1945), all of whom subscribed to a neuromaturational position regarding motor development. Gesell perhaps did more than anyone to develop a theoretical statement of this position. As Connelly (1986, p.5) describes, "In observing that development progressed through an orderly sequence of stages, Gesell believed that the sequence itself was fixed by biological factors which emerged through the evolutionary history of the species. The rate of progression through the sequence of stages was considered a function of an individual's genotype but the broad pattern itself was one typical of the species." Thus Gesell dealt with two

apparently paradoxical features of development, similarity and variation between individuals.

McGraw was among the first to link the emergence of motor skills to the emerging organization of the nervous system. Specifically, this concerns the development of cortical control over behaviour. Initially, the newborn's behaviour is seen as being primarily reflexive, due to its regulation by subcortical mechanisms. "As the cortex develops, it acquires the function of inhibition. In consequence, it commands the function of the subcortical centers, and more and more behaviour becomes voluntary." (Fitzgerald, et al., 1982, p. 107). Another principle of development that is closely tied to the changing structure of the nervous system is the notion of developmental direction. This states that motor development proceeds from the head to the feet (cephalocaudally) and from the midline of the body to its periphery (proximodistally). Thus the infant is expected to gain sequential control over the musculature of the head, neck and trunk prior to gaining control over the legs. Likewise, control of the trunk and shoulders should precede that of the wrists, hands and fingers (Gallahue, 1976).

The role of the environment in infant motor development does receive some attention in the literature. While environmental influence is not generally put forward as an

explanation for the emergence of new motor behaviours, it is often seen as having an impact on the quality and frequency of performance of certain behaviours (Gallahue, 1976). It is also probably true that environment is more influential in later infancy than in earlier infancy. As Saint-Anne Dargassies states, "from birth to three months, the maturational processes are enough to ensure the appearance of the functions found at this age; from this stage onwards, they must blend with an affective element which stimulates attention to the surrounding world; after six months, the milieu asserts itself by bringing experiences and an apprenticeship which have a part to play in the quality of acquired function." (1986, p. 12).

It is interesting to note that, in discussions of the relative importance of genetics and environment to motor development, 'environment' seems to be equated with 'learning'. If one were to expand the view of environment to include physical or chemical factors such as the supply of oxygen, it becomes clear that certain environmental conditions or events can have a profound influence on motor development. Certainly, one of the most common causes of neuromotor disability in infants is encephalopathy resulting form perinatal oxygen deprivation. Still, we might conclude that, given that certain basic environmental conditions are

met, motor development is primarily under genetic (maturational) control.

While most of the early literature on motor development focuses on a neuromaturational approach, more recent literature (e.g., Thelen, 1987) suggests alternative conceptual models for explaining the emergence of motor behaviours. One such model is the systems model of early motor development, which differs from the neuromaturational approach in the following ways.

The neuromaturational theory states that motor development is largely genetically driven, or 'hard-wired', and that the emergence of motor skills is dependent upon the degree of maturation of the central nervous system. Little importance is accorded the role of environmental factors. The systems approach, while acknowledging the key role of central nervous system maturation, also recognizes other factors as crucial to the successful performance of a motor behaviour. The infant, his environment, and the properties of the task are viewed as an integrated unit or system, and the manifest motor behaviours are viewed as the products of the interactions within this system.

Newell (1986) has identified three different categories of factors influencing behaviours within a motor system. These are:

(1) Organismic constraints - limitations on motor behaviours imposed by characteristics of the infant, such as CNS maturation, muscle strength, and biomechanical factors.

(2) Environmental constraints - environmental factors which are not related to a specific task, but which can influence performance of a motor behaviour. Gravity is the most obvious example of such a constraint, but others include temperature, noise level and lighting.

(3) Task constraints - restrictions imposed by the nature of the task or the properties of the object (e.g., the size or shape of a toy, or the evenness of a walking surface). The unique features of a task can shape an infant's motor development, and established motor behaviours can also be altered for specific tasks.

The systems approach represents a more holistic view of early motor development than does the more traditional approach. It also suggests that there may be several dimensions underlying motor performance in infants and young children rather than the strong single dimension suggested by the neuromaturational theory. It is likely that the emergence of early motor behaviours is mainly influenced by maturation, with the other dimensions playing a greater part in the refinement of specific motor skills later in childhood. In spite of some of the appealing features of systems theory, the approach that was taken in the

construction of these items was largely a neuromaturational one, particularly in reference to the assumption that the sequence of emergence of motor behaviours is predictable. Although environmental and task constraints may be very relevant with some developmental scales, the AIMS involves observational assessment, carried out in a standard clinic setting with very little use of toys or equipment, so properties of the setting and task seemed somewhat less important for this research.

The concept of individual differences in motor development has received considerable attention in the literature, even though developmental sequences of behaviour are portrayed as being reasonably invariant. The prevailing conceptualization seems to be one of regular, predictable sequences of development, superimposed on which are substantial individual differences in the rate of emergence of behaviours (therefore in the age of emergence). This pattern is certainly evident when one examines published normative scales for assessing infant development. There is a striking consistency in the ordering of motor behaviours on these scales, yet there are often differences in the typical ages for a given behaviour as well as fairly wide age bands for the emergence of a given behaviour.

Another interesting and important characteristic of early motor development is that, as new behaviours appear in

the infant's repertoire, earlier behaviours often disappear. In describing this phenomenon, Coombs and Smith (1973) state that one can think of the individual as being represented by two independent processes - one running ahead and acquiring behaviours and the other trailing behind and deleting them. The behaviours observed in an individual at any point in time, then, are those that have been acquired but not deleted. Somewhat related to this is the notion that contiguous stages or behaviour patterns are not neatly separated in the activity of the infant. Gesell and Ames (1940) point out that during the course of a day, or even of a minute, an infant may display: (a) a pattern which he has almost outgrown but reverts to for practical reasons, (b) the pattern which is most characteristic of him at his level of maturity, and (c) a pattern which is so new that he manifests it only sketchily or imperfectly.

The most obvious examples of early behaviours which are subsequently deleted are the primitive reflexes of the newborn infant, such as the rooting reflex, the palmar grasp, and the automatic walking reflex. While these reflexes are present at the time of full-term birth, they will disappear during the first few weeks or months of life. Once they are gone, they can no longer be elicited in the infant, although they may have been precursors to a later behaviour. It is possible that failure to delete a

behaviour within a specified time frame may be as indicative of motor problems as is delayed emergence of behaviours. According to Saint-Anne Dargassies (1986), one can often observe in abnormal infants a persistence of the automatic walking reflex between birth and one year of age.

In discussions of developmental sequences, the point is often made that the mere coexistence of behaviours in a sequence does not imply how they are related to one another. In a conceptual analysis of the bases for sequential order, Flavell (1972) describes five possible types of relations between successive achievements:

(1) Addition - in this relation, the later achievement does not displace but is simply added to the earlier one.

(2) Substitution - the later achievement replaces the preceding one within a sphere of activity.

(3) Modification - the antecedent achievement is transformed into the subsequent one and, therefore, should no longer be evident once the subsequent one is demonstrated.

(4) Inclusion - the earlier achievement becomes a constituent part of the later one.

(5) Mediation - the earlier achievement serves as a prerequisite step in the construction of the later one but is itself not integrated in the subsequent achievement.

In a discussion of sequential order of cognitive development, Uzgiris (1987) takes the position that it is the modification relation which characterizes true developmental sequences. No such theoretical statements have been found in the literature concerning the kinds of relations characterizing motor sequences. Still, these relationships are interesting to contemplate and are potentially quite important to the scaling of measures of motor development. For example, if the relationship between the two behaviours was one of inclusion, the first would form a part of the other; therefore, reversal of their order of attainment would be logically impossible (Flavell, 1972).

A second point made by Uzgiris concerns the integration of several achievements allowing for the emergence of a new behaviour. "If there is some substitutability in the achievements that can be integrated to form the basis for the new competence, several patterns of achievement may be compatible with the development of the higher-level competence." (Uzgiris, 1987, p.133). This idea is consistent with the principle of equifinality in development and with Waddington's model of the epigenetic landscape (Bower, 1982). The essence of this model is that there may be many different routes to the same developmental end state. The notion of different routes to some behavioral endpoint is intuitively applicable to infant motor

development and may explain why most attempts at measuring motor development have resorted to a 'developmental milestone' approach. That is, it is possible that there are relatively few behaviours that are absolutely essential steps along the road to an endpoint such as independent walking.

In summary, it appears that the literature identifies the important features of early motor development as follows:

(1) At birth, the infant possesses a number of motor responses (thus a measuring scale beginning at birth does not possess a meaningful zero reference).

(2) Early motor performance may be inflenced by a variety of factors, including maturation, environment, task and motivation.

(3) The sequence of motor development is reasonably invariant, but there are individual differences in rate.

(4) At any point in time, an infant's repertoire consists of behaviours acquired but not yet deleted.

(5) There may be several different routes leading to the same endpoint of motor behaviour, and these may all be normal.

Before concluding this section of the literature review, it seems important to briefly address the issue of age as a variable in developmental research. The normative

approach in child development relies heavily upon characteristic age descriptions, age being used as an indication of developmental status and often as a proxy measure of process (Connelly, 1986). Most criticisms of this approach center around the danger that we might come to think of age as an agent of change, rather than as just a convenient yardstick along which sequential behaviours may be anchored (Uzgiris and Hunt, 1975). Other concerns are raised by Angoff (1971) in his thorough discussion of the use of age-equivalent scores. However, as Gesell and Amatruda (1964, p.6) point out, we cannot measure development without some anchoring system since there is no absolute unit of growth. "It takes time to mature. We express the amount of time consumed by age." Wohlwill (1970, 1973) also supports the use of age, provided it is viewed simply as a dimension along which the behaviours of interest are to be studied. In this study, the continuum of chronological age was used as an anchoring system for the scaling of motor behaviour items.

Review of Scaling Models

There is a seemingly infinite number of ways in which scaling models may be classified, and a few of the most common are presented here:

(1) Dimensionality - some models locate stimuli (items) and/or persons on a unidimensional continuum, while others are capable of treating the multidimensional case.

(2) Allowance for Error - some scaling models recognize the potential for error in the data, while others do not. The former are probabilistic or stochastic models, and the latter are deterministic. In probabilistic models, any given subject response has a certain probability of occurrence, in contrast to deterministic models, in which the response is completely determined by the parameters of the subject and of the item (Torgerson, 1958).

(3) Nature of the Response - in some methods, concern is with the relationship between the stimulus (item) and the subject; in others it is with the relationship between the stimulus and the attribute.

(4) Properties of the final scale - i.e., whether ordinal, interval or ratio.

Torgerson (1958) makes the following points regarding the properties of the scale. If we are to represent an attribute, an isomorphism or one-to-one relationship must exist between the characteristics of the number system and the relations between the quantities of the attribute to be measured. The formal number system possesses the properties of order (numbers are ordered), distances (differences between numbers are ordered), and origin (the zero point).

Order is involved in all scaling methods, so the types of scales achieved are distinguished by which of the other two properties they possess. This gives rise to four types of scales: (1) ordinal, (2) ordinal with natural origin, (3) interval, and (4) ratio (Torgerson, 1958). In this study, the properties of order and distance both seem relevant to the task of scaling this set of motor development items. The meaning of distance between items is dependent upon the validity of the assumption that chronological age is an acceptable reference for motor maturation. The literature on infant motor development seems to indicate that the maturation of the central nervous system proceeds in a fairly regular, continuous fashion over time, and is directly linked to age. Thus, even though the manifest motor behaviours may emerge in somewhat discontinuous stages or patterns, the underlying maturational process that we are attempting to measure through those behaviours is probably continuous and closely mapped on chronological age.

The existence of a natural origin for motor development makes sense conceptually, but it would be difficult to get agreement as to which point in fetal development represents the true zero for motor ability. Thus, the new scale is assumed to be an ordinal or quasi-interval type of scale.

The Issue of Dimensionality. Determination of the number of dimensions underlying a data set is a central problem in measurement, and one that must be addressed prior to the application of specific methods for the scaling of items. Some models, such as multidimensional scaling and factor analysis, provide evidence for the number of underlying dimensions, while other models assume unidimensionality. Thus, the latter should be applied only after unidimensionality has been demonstrated, or should be applied to unidimensional subtests.

In regard to the measurement of developmental dimensions, Wohlwill (1973) discusses several criteria for determining when an appropriate developmental dimension has been identified. These are:

(a) Systematic shifts with age should be observed on the dimension.

(b) The dimension should have a meaningful reference to known or postulated developmental processes.

(c) The responses defining the dimension must constitute a homogeneous, unidimensional set, both within and across age levels.

(d) The dimension should be defined in general terms, sufficiently situation-independent to give a valid, stable measure of developmental status.

Following is a discussion of some scaling models that seemed most appropriate for use with the Alberta Infant Motor Scale.

Multidimensional Scaling. The several techniques known as multidimensional scaling (MDS) are characterized by their representation of the structure in data as a geometric model or picture. The objects under study (items, persons, etc.) are represented by points in the spatial model such that the significant features of the data about these objects are revealed in the relations among the points. In order to capture the complexity of the data, the points may be allowed to assume positions within a two-dimensional plane or in a space of any higher number of dimensions (Shepard, et al., 1972). Goodness-of-fit indicators, called 'stress' measures, are usually employed in multidimensional scaling to help in determining how many dimensions are appropriate to fit the data (Kruskal and Wish, 1978). The scaling of items then takes place along each of the one or more dimensions.

Multidimensional scaling procedures all make use of measures of similarity (or dissimilarity) between objects as input. However, a distinction is made between metric and nonmetric MDS. Nonmetric scaling tries only to fit the rank order of the similarities to the distances in the stimulus space, whereas classical metric scaling attempts to fit the

similarities to the distances. In general, nonmetric scaling provides better fit in low dimensionality, since it is merely trying to maintain the rank order relationships among stimuli (Schiffman, et al., 1981).

Because MDS is more flexible than some other approaches, in terms of allowing for more than one dimension, it seemed to be a reasonable starting point for data analysis in this study.

Factor Analysis. Factor analysis is a very common approach to assessing the dimensionality of a data set. In fact, McDonald (1985, p. 218) defines a unidimensional test as "a test whose items fit a latent trait or common factor model, possibly non-linear, with just one latent trait or common factor." In factor analytic models, the common factor is what the items have in common, in the sense that it explains their correlated parts. If more than one common factor is required to account for the item intercorrelations, then this would suggest that the test is not unidimensional.

McDonald (1981) has raised a concern about the application of linear factor models to binary data, because of the failure of such data to meet the assumption of linear regression of item scores on the common factor. In this study such a concern seems justified, based upon several attempts to apply different linear factor models to an

earlier subset of these data. These early findings were completely uninterpretable, and led to the decision to test a nonlinear factor analytic model on the final data set. Nonlinear factor analysis is possible using the program NOHARM, as developed by Fraser in 1983 and described by Fraser and McDonald (1988).

Item Response Models. Item Response Theory, also called Latent Trait Theory, has provided several models which can be useful in scaling. In these probabilistic models, it is assumed that a one-dimensional latent attribute (e.g., motor maturity) exists on an underlying continuum, and that the probability of passing any item increases monotonically with the levels of that ability. The parameters estimated by item response models provide for the scaling of people and of items. In a developmental scale, the person parameter represents how far the individual has progressed in the acquisition of the ability, while the item difficulty parameter indicates the position of that item on the developmental continuum (Kingma and Ten Vergert, 1985). The discrimination parameter can be thought of as an indication of whether the ability emerges slowly or abruptly over time.

Within item response theory, models are distinguished by the shape of the item characteristic curves (i.e., normal ogive or logistic) and by the number of parameters being

estimated. Models in which only item difficulties are of concern are referred to as one-parameter logistic models or Rasch models. Two-parameter logistic models are those concerned with both the item difficulty and the discrimination parameters (Crocker and Algina, 1986). Certainly the difficulty parameter is of prime importance when scaling a set of items, since it represents the scale value of each item on the ability continuum. The discrimination parameter is also important, as an indication of the abruptness of emergence of a behaviour. The additional parameter of the three-parameter model, guessing, is irrelevant for this study. With a motor development scale, it is impossible to imagine an infant accidentally performing a skill of which he is not yet capable. Thus, in applying item response theory in this study, two-parameter models appeared to be the most suitable.

Guttman Scaling. This approach was developed by Guttman in the 1940's as a method for scaling attitudes, but has been adapted for scaling sequences of achievements. It assumes that if a set of items can be ordered from lowest to highest on some ability, the scores of individuals on those items should fall into predictable patterns (Uzgiris, 1987). Anyone who has reached a certain level of competence in a sequence should demonstrate all lower levels of competence and fail to show all higher levels. The classic Guttman

model is both deterministic and cumulative. For a cumulative model to be completely appropriate for scaling motor development items, the items would have to be related in an additive way (as described in Flavell's 1972 analysis). The model makes the assumption that lower levels of achievement can be assessed even when higher levels have been attained. Since this is clearly not the case with many developmental sequences, some authors (Wohlwill, 1973; Coombs and Smith, 1973; Uzgiris, 1987) have recommended disjunctive models for such data sets.

In a disjunctive model, an individual's ability is assumed to correspond best to some point on a continuum, and he is expected to perform only a few behaviours that best fit his ability. The model requires that all demonstrated behaviours be adjacent in the developmental sequence but does not demand performance of all levels below the highest one (Uzgiris, 1987). Coombs and Smith (1973) developed a general model of scaling for disjunctive data, but did not provide any analytic procedures for handling such data. To the author's knowledge, no further work has been carried out on these models, so they were not employed in this study.

The Structural Validity Context

The application of any of the described scaling models is essentially an effort to search for the structure in the item data, in order to gather evidence as to the construct

validity of the new instrument. The lines of evidence which establish the construct validity of a test refer to its content, its internal structure, and its relation to outside variables. Loevinger (1967) calls these three components: (1) substantive validity, (2) structural validity, and (3) external validity. In the development of the Alberta Infant Motor Scale, substantive validity has been addressed through the process of item generation and consultation with content experts. External validity will be examined in part by the correlation of AIMS scores with those of other known tests. The focus of the scaling study, then, is on the issue of structural validity. The structural component of validity refers to the extent to which structural relations between test items parallel what is known about the nature of the trait being measured (Loevinger, 1967). In this case, the term 'structure' includes the number of dimensions in the data, the relative position of items, and the extent to which the scoring system reflects the nature of the underlying construct.

In addition to statistical criteria for goodness-of-fit of a given model to the data, one must apply some logical criteria to the selection of scaling models and the interpretation of results. Loevinger recommends the following questions to keep in mind in evaluating various models used in examining structural validity:

(1) Does the chosen structure correspond to what is known about non-test manifestations of the trait?

(2) Is the degree of fit quantitatively evaluated?

(3) Is the model used for selecting data? (If so, it should be re-evaluated on a new sample).

(4) Are the parameters of structure (e.g., the number of factors) uniquely determined?

With regard to the construct of infant motor maturity, both theory and clinical practice tell us that the trait should be unidimensional, that certain behaviours should appear and then disappear over time, and that not all infants will follow the same behavioural path to a given endpoint in development. These principles have been used as a guide for designing this study, for the selection of scaling models and evaluation of their fit to the data, and for the examination of possible scoring systems to be employed with the Alberta Infant Motor Scale.

Chapter 3

Methods

Sample

The subjects for this study were 506 normal infants recruited through the Edmonton Board of Health well-baby clinics and meeting the following inclusion criteria:

(1) gestational age of 38-42 weeks at the time of birth

(2) birth weight of > 2500 grams

(3) uncomplicated delivery

(4) deemed normal upon discharge from hospital

(5) no obvious abnormality at the time of assessment

The sample size was largely determined by the total number of items in the instrument, which is 58. A sample of approximately 500 was believed to be a reasonable compromise between a sufficient number of subjects to conduct a factor analysis and the costs of testing each infant.

The sample was age-stratified, by month, through the first eighteen months of life. The upper age limit of eighteen months was chosen so as to be reasonably certain of capturing the age of independent walking in all normal infants. The instrument was constructed with the intent that it would be most sensitive around the middle of the first year of life, since that is generally considered to be the optimal time to identify infants who have a motor delay

and to commence treatment programs for them. For this reason, slightly larger numbers of children were sampled in the age categories between three and twelve months than in the very young or older age categories.

The decision to use only normal infants for this part of the study arose from the need to first document the sequence and age of emergence of these motor behaviours within normal development. A set of items sequenced in this manner then becomes an appropriate reference point for assessing the instrument's sensitivity in detecting deviations from normal development in other groups of infants.

Design

Since each infant was tested only once for the scaling study, the design was a cross-sectional one. With this type of design, the order of emergence of behaviours is reflected in the mean or median ages of infants passing items to that level and not beyond, and individual differences in development are indicated by the variability in the ages. Hypotheses regarding the invariance of the order of emergence across individuals cannot be properly tested without employing a longitudinal design. Lerner (1986) and Wohlwill (1973) point out that the main criticism of cross-sectional designs for developmental research is the confounding of age with birth cohort. The concern here is

that individuals varying in age, when tested at a given point in time, represent different generations of individuals, which could vary in numerous ways besides on the variable of interest. However, with an age range of only eighteen months in the sample, this problem should be minimal. In addition, a cross-sectional design has certain advantages over a longitudinal one, such as the considerable saving of time and the avoidance of problems of attrition.

Data Collection Procedures

Data were collected over a period of fifteen months, beginning in December, 1989 and ending in March, 1991. Infants were assessed by one of six pediatric physical therapists who were experienced in infant motor assessment and trained in the administration of the AIMS. This study was carried out as part of the testing of interrater reliability, test-retest reliability and concurrent validity. This reliability and validity assessment required that some infants be tested by more than one rater and also tested on a second occasion. In addition, certain infants were simultaneously tested on the AIMS, the Peabody motor scale and the Bayley motor scale. However, for the scaling study, only one of these assessments was used, and in each case it was the initial AIMS assessment by the primary rater.

The findings of the reliability study, which involved 240 infants, indicated total test score reliability estimates above .98 for: interrater reliability (n=221), test-retest reliability (same rater; n=95) and test-retest reliability (different raters; n=138). These estimates were computed excluding infants with perfect scores, and including mean differences as a source of error. Because it was possible that the wide range of ability in the sample might account for the very high reliabilities, estimates were also obtained on subgroups of infants, grouped by three-month age intervals. Even within age groups, the reliability estimates all remained above .90.

Scoring of Items

Scoring of items was done at the time of assessment, on a five-page set of score sheets, as shown in Appendix B. For each position (prone, supine, sitting and standing), items were sequenced according to their developmental order, to the extent that we could identify their order from our earlier work with the content experts and with the data on the 97 infants used in the feasibility test.

It had been the intent, as a result of the input received from the panel of experts, to score items as pass/fail (1 or 0). However, some concern arose regarding using the simple pass/fail method during this phase of instrument construction because: (a) we were unclear as to

the exact order of items, especially for items close to each other developmentally, and (b) as the infant matures, he often loses the ability to perform earlier behaviours. These two factors presented a problem in relation to behaviours that were not observed during a testing session. Failure to observe a behaviour in a particular infant could have several possible interpretations, for example:

(a) The infant is currently capable of performing the behaviour, but simply does not do so during testing.

(b) The infant is not yet capable of that behaviour.

(c) The infant has performed the behaviour in the past, but it is no longer in his repertoire.

In order to deal with this problem, an item scoring method was used which allowed for a third category called 'assumed previous pass', to be assigned at the discretion of the physical therapist doing the assessment. The full set of scoring instructions were outlined for the therapists in the data collection manual, as follows:

> "Please score all items which you believe are at the current motor skill level of the infant. Continue to score items until you are confident that subsequent items are beyond the infant's present motor abilities. Because the items may be out of developmental order, please score several items above and below the infant's current motor level to assist us in determining the correct item sequence.
>
> Items at the infant's current level or beyond should be scored on a pass/fail (P/F) basis according to whether or not you have observed that behaviour in the infant during the testing session. In other words, the score of 'pass' (P) should only

be given to an item when the infant has actually demonstrated that behaviour.

A score of 'fail' (F) for an item should be given if:

(1) the infant does not exhibit the behaviour because he is not yet capable of it.

(2) the infant attempts the behaviour but does not perform it correctly.

(3) you believe the behaviour may be in the infant's current repertoire but you do not observe it during the testing session.

Items depicting behaviours which are no longer in the infant's repertoire, but which you believe he has previously performed, should be scored as 'assumed previous pass' (APP). When you have finished scoring the items within each position, you should have in addition to some items scored as P (1), several items scored as APP (2) and/or F (0)."

When the coding of data was complete a typical set of scores for an infant within a single position had the following pattern:

2 2 2 2 2 2 2 2 1 1 1 0 1 0 1 0 0 0 0 0 0 0

Following the examination of some descriptive statistics on the raw item scores, all items were recoded so that 'assumed previous pass' scores were treated as 'passes'. The result was a dichotomously scored (0/1) data set which was then used in all subsequent analyses. This data set is included as Appendix F.

Data Analysis

The data analyses included a descriptive accounting of the demographic characteristics of the sample and an examination of the relationship between the chronological age and the actual post-conceptional age of the infants.

Because age was believed to be so intrinsically related to the trait of interest (motor maturity), it was a logical choice as a criterion variable for assessing the fit of certain models to the data. Therefore, it was important to determine whether chronological age was suitable for this purpose or whether some adjustment should be made for gestational age at birth.

The further analyses, which included tests of dimensionality and procedures for scaling items, were carried out only on data from 479 infants who ranged in age from 0 to 15 months. Although our sample consisted of infants up to 18 months of age (n=506), we found that essentially all of the older infants had passed every item. As Wohlwill (1973, p. 112) points out "The scalability of any response matrix can be arbitrarily enhanced by ensuring a sufficiently large number of cases of subjects responding to or passing either all or none of the items, which necessarily constitute perfect scale patterns." He further indicates that this problem is most severe in the study of developmental sequences which concern only a limited portion of an age continuum. The restriction of the sample to only those infants expected to have some mixture of pass/fail scores seemed a sensible way to minimize this problem.

Tests of Dimensionality. Multidimensional scaling was employed as the primary means of assessing the

dimensionality of the data set. This was performed using a nonmetric procedure with ALSCAL (Young, Takane and Lewyckyj, 1978). The distance measure selected in the creation of the dissimilarities matrix for input into ALSCAL was the Euclidean distance for binary items. Goodness of fit indices used with multidimensional scaling were Kruskal's Stress value and the squared correlation between distances and dissimilarities. In accordance with Wohlwill's (1973) recommendation that dimensionality be tested both across and within age levels, multidimensional scaling was applied first to all data from infants 0 to 15 months of age, and then to data from several individual age groupings.

Dimensionality was also examined through nonlinear factor analysis, using the program NOHARM (Fraser, 1988). The fit of a one-factor model was assessed by examining the factor loadings, the unique variances and the residual covariances. Comparison with a two-factor model was carried out by applying the Incremental Fit Index (De Champlain and Gessaroli, 1991) to the residuals.

Some further information regarding dimensionality was gathered through the application of Guttman scalogram analysis, with its two goodness-of-fit indices, the scalability coefficient and the reproducibility coefficient.

Methods for Scaling Items. Nonmetric multidimensional scaling was the first model applied for the purposes of

scaling the items, since it allowed for the option of scaling on as many dimensions as were found to be necessary to adequately fit the data. The nonmetric approach had the advantage of producing scale values with interval level properties, while only requiring ordinal level assumptions about the relationships in the original data.

A second approach to item scaling was to examine item difficulty estimates derived according to various models. The method used first was a simple calculation of the proportion of infants passing each item, and this was obtained using the program LERTAP (Nelson, 1974). Next, a two-parameter item response model was applied to the data, using the program LOGIST (Wood, Wingersky and Lord, 1976), and the item difficulty estimates obtained. A one-parameter (Rasch) model would have been sufficient for this purpose, but attempts to perform this analysis using BICAL (Wright, 1979) on an earlier data set had proven unsuccessful, possibly because several of the items had biserial correlation estimates greater than one.

Item difficulty estimates were then obtained from the NOHARM nonlinear factor analysis results. Finally, an estimate of the age at which fifty percent of infants pass each item was derived from crosstabulations of passes/fails by age group.

The ordering of items by each of these various models was compared in order to arrive at a conclusion regarding the sequence of items. Distances between items, and their placement on the age continuum, were estimated using 'age at which fifty percent pass', in conjunction with the scale values from multidimensional scaling.

A number of correlations were calculated to assess the match between the sets of estimates arising from different scaling methods.

Item Discriminations. Item discrimination estimates were of interest as a possible way of describing the sharpness of emergence of a particular motor behaviour over time. These were derived from LOGIST, NOHARM, and TESTGRAF. TESTGRAF, which was also used to plot item characteristic curves for all 58 items, is a program developed by Ramsay (1991) for the graphical analysis of multiple choice data. The program estimates option characteristic curves through the use of a kernel smoothing technique, after examinees and responses have been ranked and assigned to certain quantiles on a standard normal distribution. The characteristic curves are then estimated by smoothing the computed relationship between the 0-1 response variable and the quantiles.

The other approach to examining the pattern of emergence of behaviours over time was to document, for each

item, the percentage of infants within each age group who passed the item. Wohlwill (1973) cautions that such group-incidence data have limited usefulness as carriers of information about the development of the individual child, although they may be of value in determining age-placement of items. He demonstrates how a group incidence function can result from a family of individual trace lines of very different shape, with the members of the family varying in the age at which the behaviour started to develop. This problem may limit the interpretation one can make from these analyses.

Determination of Scoring System. Following the scaling of items, two different scoring systems were tested on the data. One of these was a simple sum of the items passed, while the other was a score based on the highest item passed. The suitability of each scoring method was assessed by correlating the two scores with each other and with infant's age, and through discussions with the content experts.

Chapter 4

Results and Discussion

Descriptive Statistics - Infants

The final sample for the scaling study consisted of 506 infants, 285 males and 221 females. Other demographic characteristics of the sample are summarized in Table 1. Infants' gestational ages at the time of birth were very tightly clustered around 40 weeks, and the correlation between gestational age and chronological age was determined to be .998. For these reasons, chronological age appeared to be a valid indicator of post-conceptional maturity.

The actual distribution of infants across age strata is shown in Table 2. The ages of the infants were computed by the SPSSx 'Date' function, using the infant's date of birth and the date of assessment. Because this was a more exact procedure than that followed by the data collectors when selecting infants for inclusion in an age stratum, the numbers of infants are somewhat uneven across age groups. However, in accordance with our original intent, there were more infants sampled in the middle age groups (3 to 12 months) than in the very young or older groups.

Descriptive Statistics - Items

The raw score frequencies for all 58 items are given in Table 3. The most striking feature of these scores is the

frequency of 'assumed previous pass' relative to the frequency of 'pass'. This information is lost in all subsequent analyses, since these two categories are combined into a single 'pass' category. However, it seems important to bear in mind that, for a substantial number of items, relatively few 'pass' scores arose from an actual observation of that behaviour in the infant at the time of testing. For the very early items, such as Item 1 (Prone Lying 1) and Item 22 (Supine Lying 1), this is easily explained by the fact that virtually all of the infants in the sample were too old to have had any possibility of demonstrating such behaviours. More interesting, perhaps, are some of the mid-range or later items that also have a very small number of actual 'passes'. Some examples are Item 15 (Reciprocal Crawling) and Item 48 (Cruising with Rotation) which have 42 and 34 'passes', respectively. Possibly these are behaviours that exist for such a short time in the repertoires of infants that one can observe them only in infants within a very small developmental range. Thus, such behaviours were not directly observable in a large number of the infants, despite the rather large sample size.

Dimensionality of the Data Set

Multidimensional Scaling Results. Table 4 contains the results pertaining to goodness-of-fit tests from the

multidimensional scaling analyses performed using ALSCAL. These are based upon data from 479 infants (all infants up to 15 months of age). From Table 4, it is clear that a single dimension provides an excellent fit to these data, as evidenced by a stress value of .04 and RSQ of .995. Interestingly, adding a second or third dimension did nothing to improve the fit of the model to the data. Indeed, in the third decimal place, the fit actually became slightly worse with additional dimensions.

In order to determine the nature of the single dimension, the item scale values from the one-dimensional solution were examined in relation to our hypothesized order of emergence of items within each of the four assessment positions. Table 5 shows the items ordered within position, with the early items at the top of each column, according to the results of our earlier content validation work and the feasibility study. It can be seen that the scale values from the one-dimensional solution are ordered in the same manner, with a few minor exceptions. This suggests that the dimension is nothing more than a developmental sequencing one, and that the single construct underlying these data is probably gross motor maturity.

There was some concern that the one-dimensional model may have fit as well as it did due to the large variation in motor ability across the sample strata and the concomitant

large variability across the item set in terms of the level of maturity required to perform the behaviours. Nunnally (1978) cautions that it is easy to fool oneself into believing a unidimensional scale is present when one takes a set of items widely dispersed in difficulty and administers them to a very diverse population. Also, Wohlwill (1973) emphasizes the importance of testing dimensionality both within and across age levels. For these reasons, the analyses reported in Table 6 were carried out. This table gives the goodness-of-fit statistics for the two-dimensional and one-dimensional solutions when the multidimensional scaling analyses were performed on data from individual age groups. The number of infants in each age group and the number of items included in each analysis are also indicated in the table. Originally, these analyses were attempted using all 58 items, but ALSCAL encountered difficulty estimating the parameters for certain age groups. This seemed to be due to the large number of items either passed or failed by all infants within the particular age group. The analyses reported in Table 6 were performed on only those items for which scores were not constant across all infants in the specific age group.

The stress values for the one-dimensional solution ranged from .054 to .178, slightly higher than that obtained using all 479 infants. This finding probably reflects the

fact that age (i.e., maturity) was the largest contributor to the variation between item scores. Grouping the data by age category removed much of the influence of age, with the result that the variance among items within age groups appears somewhat less systematic.

Still, for the majority of age groups, these stress values for the one-dimensional solution were very close to the two-dimensional stress values, which ranged from .003 to .106. Thus, the conclusion that the data are unidimensional still seems to be a reasonable one.

Factor Analysis Results. The factor loadings and unique variance estimates obtained using NOHARM's (Fraser, 1983) one-dimensional nonlinear factor model are reported in Table 7. Six very early items (2 prone, 2 supine, 1 sitting and 1 standing) were excluded from this analysis, since the procedure was unable to converge when these extremely easy items were included. As indicated in Table 7, all 52 items had loadings of 1 or near 1 on the single dimension, and the unique variances were exceptionally small. In addition, the sum of squares of residuals for the one-factor solution was .402, and the root mean square of residuals was .0174. These very low values appear to represent further strong evidence for unidimensionality.

A new goodness-of-fit index for use with NOHARM has been suggested by De Champlain and Gessaroli (1991), and is

called the Incremental Fit Index (IFI). This involves subtracting the sum of squares (residual) of the m+1 factor solution from the sum of squares (residual) of the m factor solution, then dividing by SS (residual) for m factors. In the case of a one-factor versus a two-factor solution, the IFI calculates the proportion of residual covariance from the one-factor solution that can be accounted for by a second factor. If the second factor is important in explaining the structure of the data, then the IFI should be large.

In order to calculate the IFI for this data set, a two-dimensional exploratory nonlinear factor analysis was performed using NOHARM. The value for sum of squares of residuals was exactly the same as that obtained with a one-factor solution (.402), resulting in an IFI of zero. Thus it seems clear that no second dimension is required to account for the data.

Guttman Scaling Results. The Guttman scalogram analysis was less helpful in determining dimensionality of the data set, since the available program was only capable of handling thirty items at one time. The items we believed to be the earliest twenty-nine items were analyzed together, producing a Scalability Coefficient of .74 and a Reproducibility Coefficient of .91. Similar findings were observed for the later 29 items, with a Scalability

Coefficient of .76 and a Reproducibility Coefficient of .93. These values seem to provide at least some additional support for the unidimensionality of the two subsets of items.

Scaling of the Items

Table 8 contains the results of the application of four different approaches to determining the sequence of the 58 items on the continuum for normal motor development. The most straightforward of these approaches was to simply look at the proportion of infants passing each item. These proportions were obtained using LERTAP, and are reported in the first column of Table 8. The order that one would assign to each item based upon proportion passing is included in brackets, with '1' representing the earliest item and '58' representing the latest item.

The item difficulty parameters as estimated by a two-parameter Item Response Model (LOGIST) are reported adjacent to the LERTAP findings, in Table 8. It can be seen that the ordering of items by the LOGIST difficulties was identical to that obtained with LERTAP.

The NOHARM item difficulty estimates and the scale values from multidimensional scaling (ALSCAL), as reported in Table 8, also lead to the same ordering of items, although the difficulties of the earliest six items could not be estimated by NOHARM.

Further evidence as to the sequencing of items was gathered by examining the percentage of infants passing each item within each of the various age groups, split by completed month of age. Table 9 gives an estimated age placement for each item, based upon the age at which fifty percent of infants would be expected to pass the item. Values in the table were determined by linear interpolation between the two ages which bounded the fifty percent pass rate. These age estimates correlate very strongly (.97 to .99) with the item difficulties and the item scale values previously reported, and these correlations are shown in Table 10. It is clear that the ordering of items is very similar, regardless of which scaling approach is taken, and this provides a high degree of confidence in the validity of the item sequence.

With regard to determining the distances between items on the age scale, the estimated age at which fifty percent pass has an obvious appeal, since it is directly anchored to chronological age. However, the confidence one can place in these estimates is somewhat reduced because they were determined from group incidence data. If infants within each age grouping were distributed evenly along the one-month age interval, then the group incidence data would seem more valid. Unfortunately, with this sample that is not the case, since within some age groups infants cluster towards

one end of the one-month interval. This reflects the fact that infants typically are brought to the health units at specific ages (2 months, 4 months, 6 months, etc.), to receive immunization. So, for example, our six month age group, which should include infants between six and seven months, has a higher number of infants near the bottom of that interval than near the top. This problem could lead us to overestimate the difficulty of certain items in that region, i.e., place the item slighter later on the age continuum than it actually should be placed.

For the above reasons, an age scaling technique was sought which would make use of raw item scores on all infants, rather than on grouped data. The multidimensional scale values, transformed onto a chronological age scale, seemed the most appropriate for this purpose, and these estimates are reported in Table 9, alongside the age estimates based upon the group incidence data. The procedure followed in rescaling the MDS values was to add the constant 1.68 and then multiply by the constant 3.6, converting them to a scale beginning at age zero and ending at 12.73 months, which were the lowest and highest estimates from the grouped data.

Depending upon the scoring system used with the final scale, the exact placement of items on the chronological age continuum may not be important (discussed later under

'Scoring System'). However, if such placement is desired, the rescaled multidimensional scale values are recommended for this purpose, since they seem to more accurately reflect the full information within the data set.

Item Discrimination Estimates

The estimation of item discrimination parameters for these items proved to be somewhat problematic. The findings are summarized in Table 11, and include estimates obtained under LOGIST, NOHARM, and TESTGRAF. Generally, it appears that these are very highly discriminating items, at least in comparison to achievement-type data, for which most of these measurement models were derived. Although it is theoretically possible for discrimination values to vary from negative to positive infinity, in practice they usually take on values only between zero and two.

As shown in Table 11, LOGIST simply set all the values to 5.05, indicating that the program was unable to properly estimate the parameter and placed it at a default ceiling value. By other criteria, though, the two-parameter logistic model seemed to provide a good fit to the data. For example, the difficulty estimates corresponded to those estimated by other methods, and the ability estimates (theta values) obtained with LOGIST showed a .95 correlation with infant's chronological age.

NOHARM produced discrimination estimates with some variability between items, but many of the values are so huge they seem to escape interpretability. The most reasonable estimates arose from TESTGRAF, although even many of these were larger than one would normally expect.

TESTGRAF was employed to plot characteristic curves for all items. These item characteristic curves are included in Appendix C. On each of these graphs, the solid ascending line represents the curve for the 'pass' option, and the crosshatching on this line is an indication of the standard error at each point. The dotted descending line represents the characteristic curve for the 'fail' option. The patterns evident in these plots are very consistent with the data concerning the proportion of infants passing within the various age groups.

Scoring System

Following the determination of item sequence, infants' test scores were derived according to two different scoring systems:

(1) 'Pass' scores were summed for each infant, giving a total number of items passed (range of 1 to 58).

(2) Items were reordered, based upon their MDS scale values, and a score given to each infant corresponding to the position (1 to 58) of the highest item passed.

Each of these scoring systems had a certain appeal to the members of the research team. The 'number of items passed' is the system used with most well-known motor scales, and is very straightforward in its calculation. It also lends itself quite easily to the construction of norms tables, since typical total scores and ranges (or percentiles) can be reported for various age groups of infants. A total score or percentile rank is generally considered to be less problematic than certain other types of scores such as age-equivalents, particularly when reporting an infant's performance to the parents.

The 'highest item passed' system possesses the same advantages with regard to calculation, norming and reporting. However, it has the added appeal of being a compensatory model for scoring, since it gives an infant credit for the level at which he is currently capable of performing, regardless of which behaviours he performed or did not perform previously. Thus, the use of this system would allow us to avoid the situation of awarding 'pass' scores for behaviours not actually observed but assumed to have been performed at an earlier time. It is also consistent with the opinion of many experts in infant motor assessment that it is the endpoint that is important in determining an infant's motor ability, rather than the means by which he arrived at that endpoint.

In order to evaluate the two scoring systems quantitatively, both types of scores were computed for all 506 infants. The correlation between the two scores was found to be .99, indicating a high degree of consistency across these two systems, when applied to data from normal infants. In addition, infant's age had a strong relationship both with 'number of items passed' (r=.95) and 'highest item passed' (r=.94). These findings provided no real basis for choosing one scoring method over the other.

In discussing these results with the content experts, it was recognized that the consistency observed between the scoring methods might not hold when the instrument was applied to high-risk or abnormal infants. Specifically, it was felt that abnormal infants might pass certain items but be incapable of performing some earlier behaviours, particularly in a different postural position. This could give such an infant a higher score than was appropriate, if scoring was based upon highest item passed. For this reason, it was decided that the total number of items passed was a more reasonable scoring system to retain, at least until various systems could be tested on data from abnormal and high-risk infants.

With a scoring system based upon the number of items passed, the question of precise age placement of items

becomes less important, and only the correct sequencing of items is of concern.

Chapter 5

Conclusions and Recommendations

This study involved an extensive examination of the scale properties of the fifty-eight items in the Alberta Infant Motor Scale, on a sample of 506 normal Edmonton infants. The findings seem conclusive, with regard to the structure of the observed data.

It is clear that the AIMS is a unidimensional test, at least when applied to normal infants, and that the single dimension underlying the item scores is gross motor maturity. However, it is quite possible that other dimensions will emerge when subsequent data are gathered from abnormal and high-risk infants. For example, future data could suggest that the instrument should be scored as four separate subscales, corresponding to the four postural positions (prone, supine, sitting and standing), since abnormal infants sometimes perform quite differently across positions. Therefore, it is recommended that tests of dimensionality be carried out on all future validation and norming data. Nonmetric multidimensional scaling (ALSCAL) would seem to be the most useful model for this purpose, because it requires few assumptions about the data and because it produces well-known goodness-of-fit indices for any number of dimensions. It is further recommended that,

for subsequent data sets, dimensionality be examined both across and within age levels, in accordance with Wohlwill's (1973) suggestion.

With regard to the developmental sequencing of the items, all of the scaling models applied to these data suggested virtually the same sequence. Therefore, practical considerations such as ease of application and interpretability seem to be the only criteria for selecting one model over another. Again, nonmetric multidimensional scaling (ALSCAL) is recommended, for the following reasons:

(1) It can produce interval level information about the distances between items, while requiring only ordinal level assumptions about the relationships in the data.

(2) It can handle a large number of cases and variables at a time.

(3) It tests the fit of various numbers of dimensions to the data, and at the same time estimates item scale values for as many dimensions as are currently in the solution.

In summary, the findings of this study have supported earlier work regarding the construct validity of the Alberta Infant Motor Scale. All of the evidence gathered to date indicates that, at least when applied to normal infants, this is a unidimensional test of gross motor maturity. Further, there is convincing evidence of the scalability of

the fifty-eight items along this single continuum. It is recommended that items on the AIMS should be re-sequenced based upon the findings reported here, and that this should be carried out prior to the collection of normative data. Revised score sheets, indicating the recommended sequence of items, are included in Appendix E.

Certain tentative recommendations can also be made regarding the reporting of infants' scores. Following the collection of age- and sex-related normative data, these data can be most easily summarized into norms graphs which identify AIMS scores at various percentiles. An example of such a graph is given in Appendix D. It is recommended that these be constructed for males and females separately, and that an option be provided for the graphing of scores on four subscales corresponding to the four postural positions, in addition to the total test score. This reporting method would seem to be a simple and concise way of documenting an infant's motor profile at a given time and over time, and of evaluating performance against that of the appropriate reference group.

References

Anastasi, A. (1988). Psychological Testing. (6th Edition). New York: Macmillan.

Andrich, D. (1985). An elaboration of Guttman scaling with Rasch models for measurement. In N.B. Tuma (Ed.). Sociological Methodology 1985. San Francisco: Jossey-Bass .

Angoff, W.H. (1971). Scales, norms and equivalent scores. In R.L. Thorndike (Ed.). Educational Measurement (2nd Edition). Washington, D.C.: American Council on Education, 508-600.

Baird, J.C. & Noma, E. (1978). Fundamentals of Scaling and Psychophysics. New York: Wiley & Sons.

Bayley N. (1969) Bayley Scales of Infant Development. Berkeley: Institute of Human Development, University of California.

Bohrnstedt, G.W., & Borgatta, E.F. (1981). Social Measurement: Current Issues. Beverly Hills: Sage.

Bower, T.G.R. (1982). Development in Infancy. (2nd. Edition). San Francisco: W.H. Freeman.

Connelly, K.J. (1986). A perspective on motor development. In M.G. Wade & H.T.A. Whiting (Eds.), Motor Development in Children: Aspects of Coordination and Control. Boston: Martinus Nijhoff, 3-21.

Coombs, C.H. (1964). A Theory of Data. New York: Wiley & Sons.

Coombs, C.H. & Smith, J.E.K. (1973). On the detection of structure in attitudes and developmental processes. Psychological Review, 80(5), 337-351.

Corbin, C.B. (1973). A Textbook of Motor Development. Dubuque, Iowa: Wm. C. Brown Company.

Crocker, L. & Algina, J. (1986). Introduction to Classical and Modern Test Theory. New York: Holt, Rinehart & Winston.

Davies, P.M., & Coxon, A.P.M. (1982). Key Texts in Multidimensional Scaling. London, England: Heinemann Educational Books.

Davison, M. (1980). A psychological scaling model for testing order hypotheses. The British Journal of Mathematical and Statistical Psychology, 33, 123-141.

De Champlain, A.D. & Gessaroli, M.E. (1991). Assessing Test Dimensionality Using an Index Based on Nonlinear Factor Analysis. Paper presented at the American Educational Research Association Meeting, Chicago, Illinois.

Dunn-Rankin, P. (1983). Scaling Methods. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Embretson, S.E. (1985). Test Design: Developments in Psychology and Psychometrics. Toronto: Academic Press.

Fitzgerald, H.E., Strommen, E.A., & McKinney, J.P. (1982). Developmental Psychology: The Infant and Young Child. Homewood, Illinois: The Dorsey Press.

Flavell, J.H. (1972). An analysis of cognitive-developmental sequences. Genetic Psychology Monographs, 86, 279-350.

Folio, M.R., Flewell, R.R. (1983). Peabody Developmental Motor Scales and Activity Cards. Allen, Texas: DIM Teaching Resources.

Fraser, C. & McDonald, R.P. (1988). NOHARM: Least squares item factor analysis. Multivariate Behavioral Research. 23, 267-269.

Gallahue, D.L. (1976). Motor Development and Movement Experiences for Young Children. New York: John Wiley & Sons.

Gesell, A. (1940). The First Five Years of Life. New York: Harper & Brothers.

Gesell, A. & Ames, L.B. (1940). The ontogenetic organization of prone behaviour in human infancy. The Journal of Genetic Psychology, 56, 247-263.

Gesell, A., & Amatruda, C.S. (1964). Developmental Diagnosis. New York: Harper & Row.

Griffiths, R. (1954). The Abilities of Babies. London: University of London Press.

Gulliksen, H. (1967). Mathematical solutions for psychological problems. In D.N. Jackson & S. Messick (Eds.), Problems in Human Assessment. New York: McGraw-Hill, 3-23.

Hattie, J. (1984). An empirical study of various indices for determining unidimensionality. Multivariate Behavioral Research, 19, 49-78.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. Applied Psychological Measurement, 9(2), 139-164.

Hulin, C.L., Drasgow, F., & Parsons, C.K. Item Resonse Theory: Application to Psychological Measurement. Homewood, Illinois: Dow Jones-Irwin.

Keogh, J., & Sugden, D. (1985). Movement Skill Development. New York: Macmillan.

Kingma, J., & Ten Vergert, E.M. (1985). A nonparametric scale analysis of the development of conservation. Applied Psychological Measurement. 9(4), 375-387.

Kruskal, J.B., & Wish, M. (1978). Multidimensional Scaling. Beverly Hills: Sage.

Lerner, R.M. (1986). Concepts and Theories of Human Development. New York: Random House.

Loevinger, J. (1967). Objective tests as instruments of psychological theory. In D.N. Jackson and S. Messick (Eds.), Problems in Human Assessment. New York: McGraw-Hill, 78-120.

McDonald, R.P. (1981). The dimensionality of tests and items. The British Journal of Mathematical and Statistical Psychology, 34, 100-117.

McDonald, R.P. (1982). Some alternative approaches to the improvement of measurement in education and psychology: Fitting latent trait models. In D. Spearrit (Ed.), The Improvement of Measurement in Education and Psychology. Hawthorne, Australia: Australian Council for Educational Research.

McDonald, R.P. (1985). Factor Analysis and Related Methods. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

McGraw, M.B. (1945). The Neuro-muscular Maturation of the Human Infant. New York: Hafner.

Nelson, L.R. (1974). Guide to LERTAP Use and Design [computer program manual]. Dunedin, New Zealand: University of Otago, Education Department.

Newell, K.M. (1986). Constraints on the development of coordination. In M.G. Wade & H.T.A. Whiting (Eds.), Motor Development in Children: Aspects of Coordination and Control. Boston: Martinus Nijhoff, 341-360.

Nunnally, J.C. (1978). Psychometric Theory. New York: McGraw-Hill.

Ramsay, J.O. (1991). TESTGRAF: A Program for the Graphical Analysis of Multiple Choice Items. McGill University (unpublished manual).

Reckase, M.D. (1984). Scaling techniques. In G. Goldstein & M. Hersen (Eds.), Handbook of Psychological Assessment. Toronto: Pergamon Press, 38-53.

Saint-Anne Dargassies, S. (1986). The Neuro-motor and Psycho-affective development of the Infant. New York: Elsevier.

Sciffman, S.S., Reynolds, M.L., Young, F.W. (1981). Introduction to Multidimensional Scaling. New York: Academic Press.

Shepard, R.N., Romney, A.K., & Nerlove, S.B. (1972). Multidimensional Scaling (Vol. 1). New York: Seminar Press.

Shirley, M.M. (1931). The first two years: a study of twenty-five babies. Postural and Locomotor Development. Minneapolis: University of Minnesota Press.

Thelen, E. (1987). The role of motor development in developmental psychology: a view of the past and an agenda for the future. In: N. Eisenberg (Ed.), Contemporary Topics in Developmental Psychology. New York: Wiley, 3-33.

Thorndike, R.L. (1982). Educational measurement - theory and practice. In D. Spearritt (Ed.), The Improvement of Measurement in Education and Psychology. Hawthorne, Australia: Australian Council for Educational Research.

Thurstone, L.L. (1926). The mental age concept. Psychological Review, 33, 268-278.

Torgerson, W.S. (1958). Theory and Methods of Scaling. New York: Wiley & Sons.

Tucker, L.R. (1983). Searching for structure in binary data. In H. Wainer & S. Messick (Eds.), Principles of Modern Psychological Measurement. Hillsdale, N.J.: Lawrence Erlbaum Associates, 215-283.

Uzgiris, I.C. (1987). The study of sequential order in cognitive development. In I.C. Uzgiris & J.M. Hunt (Eds.), Infant Performance and Experience: New Findings with the Ordinal Scales. Chicago: University of Illinois Press, 131-167.

Van den Daile, L. (1969). Qualitative models in developmental analysis. Development Psychology, 1, 303-310.

Van der Ven, A.H. (1980). Introduction to Scaling. New York: Wiley & Sons.

Williams, H. (1983). Perceptual and Motor Development. Englewood Cliffs, N.J.: Prentice-Hall.

Wohlwill, J.F. (1970). The age variable in psychological research. Psychological Review, 77(1), 49-64.

Wohlwill, J.F. (1973). The Study of Behavioral Development. New York: Academic Press.

Wolanski, N. (1973). A new method for the evaluation of motor development of infants. Polish Psychological Bulletin, 4, 43-53.

Wood, R.L., Wingersky, M.S., Lord, F.M. (1976). LOGIST: a computer program for estimating examinee ability and item characteristic curve parameters. Research Memorandum 76-6. Princeton, N.J.: Educational Testing Service.

Wright, B.D., Stone, M. (1979). Best Test Design. Chicago: MESA Press.

Young, F.W., Takane, Y., Lewyckyj, Y. (1978). ALSCAL: a nonmetric multidimensional scaling program with several differences options. Behavioral Research Methods and Instrumentation. 10: 451-453.

Table 1

Demographic Characteristics of the Sample

Variable	Category	No. of Infants	Percent
Sex	Male	285	56.3
	Female	221	43.7
Gestational Age at Birth	37 Weeks	25	4.9
	38 Weeks	65	12.8
	39 Weeks	92	18.2
	40 Weeks	239	47.2
	41 Weeks	68	13.4
	42 Weeks	17	3.4
Presentation	Vertex	460	90.9
	Breech	11	2.2
	Other	3	0.6
	Unknown	32	6.3
Type of Delivery	Vaginal	448	88.5
	C-Section	52	10.3
	Unknown	6	1.2

Table 2

Age Grouping of Infants

Age in Months	Age Group	No. of Infants	Percent
0 to < 1	0	12	2.4
1 to < 2	1	12	2.4
2 to < 3	2	38	7.5
3 to < 4	3	39	7.7
4 to < 5	4	40	7.9
5 to < 6	5	33	6.5
6 to < 7	6	46	9.1
7 to < 8	7	37	7.3
8 to < 9	8	34	6.7
9 to < 10	9	41	8.1
10 to < 11	10	36	7.1
11 to < 12	11	28	5.5
12 to < 13	12	35	6.9
13 to < 14	13	25	4.9
14 to < 15	14	23	4.5
15 to < 16	15	12	2.4
16 to < 17	16	5	1.0
17 to < 18	17	10	2.0

Table 3

Score Frequencies for Items

Item No.	Item Label	Assumed Previous Pass	Pass	Fail
1	P8 - Prone Lying 1	482	23	1
2	P15 - Prone Lying 2	459	29	18
3	P11 - Prone Prop	421	45	40
4	P19 - Forearm Support 1	374	56	76
5	P28 - Prone Mobility	356	57	93
6	P13 - Forearm Support 2	304	84	118
7	P6 - Rolling Prone to Supine	314	40	152
8	P20 - Extended Forearm Support	257	103	146
9	P16 - Swimming	268	66	172
10	P26 - Reaching from Forearm Support	234	98	174
11	P24 - Rolling Pr to Sup with Rotation	238	52	216
12	P5 - Pivoting	225	81	200
13	P25 - 4 Point Kneeling 1	219	56	231
14	P27 - Propped Sidelying	207	53	246
15	P1 - Reciprocal Crawling	217	42	247
16	P9 - 4 Point Kneeling to Sit	106	133	267
17	P17 - Reciprocal Creeping	175	61	270
18	P18 - 4 Point Kneeling 2	112	93	301
19	P29 - Modified 4 Point Kneeling	81	122	303
20	P23 - Reach from Extended Arm Support	66	150	290
21	P21 - Reciprocal Creeping with Rotation	49	122	335
22	Sup8 - Supine Lying 1	487	19	0
23	Sup11 - Supine Lying 2	452	48	6
24	Sup13 - Supine Lying 3	430	46	30
25	Sup3 - Supine Lying 4	333	102	71
26	Sup1 - Hands to Knees	313	81	112
27	Sup2 - Active Extension	289	82	135
28	Sup5 - Hands to Feet	256	104	146
29	Sup10 - Rolling Sup to Pr w/o Rotation	280	46	180

Table 3 Continued

Score Frequencies for Items

Item No.	Item Label	Assumed Previous Pass	Pass	Fail
30	Sup6 - Rolling Sup to Pr with Rotation	66	214	226
31	Sit13 - Sitting with Support	410	90	6
32	Sit19 - Pull to Sit	297	113	96
33	Sit18 - Sitting with Propped Arms	347	68	91
34	Sit1 - Unsustained Sitting (U.S.)	333	35	138
35	Sit22 - Sitting with Arm Support	319	27	160
36	Sit2 - U. S. w/o Arm Support	300	36	170
37	Sit23 - Weight Shifting in U. S.	290	30	186
38	Sit25 - Sitting to Prone	230	36	240
39	Sit5 - Sitting w/o Arm Support	202	95	209
40	Sit24 - Reach with Rotation in Sitting	78	190	238
41	Sit12 - Sitting to 4 Point Kneeling	26	209	271
42	Sit17 - Sitting w/o Arm Support	5	188	313
43	St8 - Supported Standing 1	465	39	2
44	St11 - Supported Standing 2	352	124	30
45	St27 - Supported Standing 3	216	138	152
46	St13 - Pull to Stand with Support	188	36	282
47	St22 - Pull to Stand, Stand w/o Support	169	49	288
48	St5 - Cruising w/o Rotation	171	34	301
49	St12 - Support Standing with Rotation	105	101	300
50	St1 - Half Kneeling	97	107	302
51	St9 - Contr'd Lowering Standing	98	95	313
52	St2 - Cruising with Rotation	111	64	331
53	St17 - Stands Alone	96	44	366
54	St21 - Early Stepping	94	26	386
55	St28 - Standing from Quadruped	31	68	407
56	St4 - Standing from Modified Squat	2	100	404
57	St10 - Walks Alone	0	97	409
58	St18 - Squat	1	83	422

Table 4

Multidimensional Scaling of 58 Items (n=479 infants)

Dimensions	Stress Value	RSQ
3	0.047	0.991
2	0.045	0.993
1	0.040	0.995

Table 5

Item Scale Values from MDS (One Dimensional Solution)

Prone Items		Supine Items		Sitting Items		Standing Items	
Item Name	Scale Value	Item Name	Scale Value	Item Name	Scale Value	Item Name	Scale Value
P8	-1.68	Sup8	-1.69	Sit13	-1.64	St8	-1.68
P15	-1.54	Sup11	-1.64	Sit19	-0.89	St11	-1.45
P11	-1.35	Sup13	-1.45	Sit18	-0.93	St27	-0.41
P19	-1.05	Sup3	-1.10	Sit1	-0.53	St13	0.70
P28	-0.91	Sup1	-0.75	Sit22	-0.34	St22	0.75
P13	-0.69	Sup2	-0.56	Sit2	-0.25	St5	0.86
P6	-0.41	Sup5	-0.47	Sit23	-0.12	St12	0.85
P20	-0.46	Sup10	-0.17	Sit25	0.34	St1	0.86
P16	-0.24	Sup6	0.22	Sit5	0.08	St9	0.96
P26	-0.22			Sit24	0.33	St2	1.11
P24	0.14			Sit12	0.60	St17	1.41
P5	0.00			Sit17	0.96	St21	1.56
P25	0.27					St28	1.73
P27	0.39					St4	1.71
P1	0.40					St10	1.75
P9	0.56					St18	1.86
P17	0.60						
P18	0.86						
P29	0.87						
P23	0.76						
P21	1.14						

Table 6

Multidimensional Scaling by Group

Age Group	Number of Infants	Number of Items	Two Dimensional		One Dimensional	
			Stress	Rsq	Stress	Rsq
2	38	13	.051	.990	.062	.989
3	39	20	.102	.965	.166	.931
4	40	28	.106	.967	.178	.925
5	33	34	.088	.974	.135	.949
6	46	38	.072	.982	.117	.962
7	37	37	.060	.987	.112	.963
8	34	34	.082	.997	.095	.974
9	41	33	.078	.984	.075	.987
10	36	33	.050	.994	.054	.994
11	28	20	.072	.988	.069	.990
12	35	27	.054	.995	.095	.985
13	25	12	.058	.986	.075	.981
14	23	11	.003	1.000	.148	.947

Table 7

Factor Analysis (NOHARM) Results

Item No.	Factor Loading	Unique Variance	Item No.	Factor Loading	Unique Variance
1	****	****	30	1.000	0.000
2	****	****	31	****	****
3	0.998	0.005	32	0.999	0.001
4	0.999	0.001	33	0.999	0.001
5	1.000	0.000	34	1.000	0.000
6	1.000	0.000	35	1.000	0.000
7	1.000	0.000	36	1.000	0.000
8	1.000	0.000	37	1.000	0.000
9	1.000	0.000	38	1.000	0.000
10	1.000	0.000	39	1.000	0.000
11	1.000	0.000	40	1.000	0.000
12	1.000	0.000	41	1.000	0.000
13	1.000	0.000	42	1.000	0.000
14	1.000	0.000	43	****	****
15	1.000	0.000	44	0.972	0.056
16	1.000	0.000	45	1.000	0.000
17	1.000	0.000	46	1.000	0.000
18	1.000	0.000	47	1.000	0.000
19	1.000	0.000	48	1.000	0.000
20	1.000	0.000	49	1.000	0.000
21	1.000	0.000	50	1.000	0.000
22	****	****	51	1.000	0.000
23	****	****	52	1.000	0.000
24	0.993	0.014	53	1.000	0.000
25	0.999	0.001	54	0.999	0.001
26	1.000	0.000	55	0.999	0.001
27	1.000	0.000	56	0.999	0.001
28	1.000	0.000	57	0.999	0.001
29	1.000	0.000	58	0.999	0.001

Table 8

Ordering of Items by Various Models

Item No.	Lertap Proportion Passing	Logist Item Difficulty	Noharm Item Difficulty	MDS Scale Values
1	99.8 (2)	-2.02 (2)	****	-1.68 (2)
2	96.2 (6)	-1.52 (6)	****	-1.54 (6)
3	91.6 (9)	-1.24 (9)	-1.39 (9)	-1.35 (9)
4	84.1 (11)	-0.89 (11)	-1.00 (11)	-1.05 (11)
5	80.6 (13)	-0.75 (13)	-0.86 (13)	-0.91 (13)
6	75.4 (16)	-0.60 (16)	-0.69 (16)	-0.69 (16)
7	68.3 (21/22)	-0.41 (21/22)	-0.48 (21/22)	-0.41 (21/22)
8	69.5 (19/20)	-0.44 (19/20)	-0.51 (19/20)	-0.46 (20)
9	64.1 (25)	-0.31 (25)	-0.36 (25)	-0.24 (25)
10	63.7 (26)	-0.30 (26)	-0.35 (26)	-0.22 (26)
11	54.9 (31)	-0.08 (31)	-0.12 (31)	0.14 (31)
12	58.2 (29)	-0.16 (29)	-0.21 (29)	0.00 (29)
13	51.8 (33)	0.01 (33)	-0.05 (33)	0.27 (33)
14	48.6 (36)	0.11 (36)	0.03 (36)	0.39 (36)
15	48.4 (37)	0.12 (37)	0.04 (37)	0.40 (37)
16	44.3 (38)	0.28 (38)	0.14 (38)	0.56 (38)
17	43.6 (39)	0.31 (39)	0.16 (39)	0.59 (39)
18	37.2 (45/46)	0.57 (45/46)	0.33 (45/46)	0.86 (45/46)
19	36.7 (48)	0.59 (48)	0.34 (48)	0.87 (48)
20	39.5 (43)	0.48 (43)	0.27 (43)	0.76 (43)
21	30.3 (52)	0.85 (52)	0.52 (52)	1.14 (52)
22	100.0 (1)	**** (1)	****	-1.69 (1)
23	98.7 (4/5)	-1.76 (4/5)	****	-1.64 (4/5)
24	93.7 (7/8)	-1.35 (7/8)	-1.54 (8)	-1.44 (8)
25	85.2 (10)	-0.93 (10)	-1.05 (10)	-1.10 (10)
26	76.6 (15)	-0.64 (15)	-0.73 (15)	-0.75 (15)
27	71.8 (17)	-0.50 (17)	-0.58 (17)	-0.56 (17)
28	69.5 (19/20)	-0.44 (19/20)	-0.51 (19/20)	-0.47 (19)
29	62.4 (27)	-0.27 (27)	-0.32 (27)	-0.17 (27)

Values in brackets indicate item order.
Items 1 to 21 are Prone Items;
Items 22 to 30 are Supine Items.

Table 8 Continued

Ordering of Items by Various Models

Item No.	Lertap Proportion Passing	Logist Item Difficulty	Noharm Item Difficulty	MDS Scale Values
30	52.8 (32)	-0.02 (32)	-0.07 (32)	0.22 (32)
31	98.7 (4/5)	-1.76 (4/5)	****	-1.64 (4/5)
32	80.0 (14)	-0.73 (14)	-0.84 (14)	-0.88 (14)
33	81.0 (12)	-0.77 (12)	-0.88 (12)	-0.92 (12)
34	71.2 (18)	-0.49 (18)	-0.56 (18)	-0.53 (18)
35	66.6 (23)	-0.37 (23)	-0.43 (23)	-0.34 (23)
36	64.5 (24)	-0.32 (24)	-0.37 (24)	-0.25 (24)
37	61.2 (28)	-0.24 (28)	-0.28 (28)	-0.12 (28)
38	49.9 (35)	0.07 (35)	0.00 (35)	0.35 (35)
39	56.4 (30)	-0.12 (30)	-0.16 (30)	0.08 (30)
40	50.3 (34)	0.06 (34)	-0.01 (34)	0.33 (34)
41	43.4 (40)	0.32 (40)	0.17 (40)	0.60 (40)
42	34.7 (49/50)	0.67 (49/50)	0.40 (49/50)	0.96 (49/50)
43	99.6 (3)	-1.93 (3)	****	-1.67 (3)
44	93.7 (7/8)	-1.35 (7/8)	-1.58 (7/8)	-1.45 (7/8)
45	68.3(21/22)	-0.41 (21/22)	-0.48 (21/22)	-0.41 (21/22)
46	41.1 (41)	0.41 (41)	0.22 (41)	0.69 (41)
47	39.9 (42)	0.46 (42)	0.26 (42)	0,75 (42)
48	37.2 (45/46)	0.57 (45/46)	0.33 (45/46)	0.86 (45/46)
49	37.4 (44)	0.56 (44)	0.32 (44)	0.85 (44)
50	37.0 (47)	0.58 (47)	0.33 (47)	0.86 (47)
51	34.7 (49/50)	0.67 (49/50)	0.40 (49/50)	0.96 (49/50)
52	30.9 (51)	0.82 (51)	0.50 (51)	1.11 (51)
53	23.6 (53)	1.09 (53)	0.72 (53)	1.40 (53)
54	19.6 (54)	1.25 (54)	0.86 (54)	1.56 (54)
55	15.4 (55)	1.48 (56)	1.02 (56)	1.73 (56)
56	15.9 (56)	1.45 (55)	1.00 (55)	1.71 (55)
57	14.8 (57)	1.52 (57)	1.05 (57)	1.75 (57)
58	12.1 (58)	1.71 (58)	1.17 (58)	1.86 (58)

Values in brackets indicate item order.
Items 31 to 42 are Sitting Items;
Items 43 to 58 are Standing Items.

Table 9

Age Placement of Items

Item No.	Age at 50% Passing	Rescaled MDS Values	Item No.	Age at 50% Passing	Rescaled MDS Values
1 (P8)	****	0.04	30 (Sup6)	7.09	6.83
2 (P15)	0.93	0.54	31 (Sit13)	<.50	0.17
3 (P11)	1.50	1.21	32 (Sit19)	3.42	2.88
4 (P19)	0.93	2.29	33 (Sit18)	3.47	2.73
5 (P28)	3.25	2.81	34 (Sit1)	4.93	4.16
6 (P13)	4.60	3.57	35 (Sit22)	5.29	4.84
7 (P6)	5.06	4.59	36 (Sit2)	6.29	5.15
8 (P20)	5.03	4.40	37 (Sit23)	6.25	5.62
9 (P16)	5.45	5.20	38 (Sit25)	7.25	7.29
10 (P26)	5.74	5.26	39 (Sit5)	6.89	6.34
11 (P24)	6.88	6.54	40 (Sit24)	7.40	7.23
12 (P5)	6.50	6.06	41 (Sit12)	8.50	8.22
13 (P25)	7.26	7.03	42 (Sit17)	9.19	9.50
14 (P27)	7.72	7.47	43 (St8)	****	0.05
15 (P1)	7.76	7.49	44 (St11)	1.20	0.87
16 (P9)	8.40	8.08	45 (St27)	5.01	4.58
17 (P17)	8.56	8.19	46 (St13)	8.75	8.55
18 (P18)	9.06	9.13	47 (St22)	8.89	8.73
19 (P29)	9.07	9.18	48 (St5)	9.03	9.14
20 (P23)	8.83	8.78	49 (St12)	9.09	9.10
21 (P21)	10.00	10.16	50 (St1)	9.03	9.16
22 (Sup8)	****	0.00	51 (St9)	9.40	9.50
23 (Sup11)	<.50	0.17	52 (St2)	9.77	10.03
24 (Sup13)	1.81	0.88	53 (St17)	11.16	11.10
25 (Sup3)	3.08	2.13	54 (St21)	11.81	11.66
26 (Sup1)	4.16	3.36	55 (St28)	12.35	12.26
27 (Sup2)	4.78	4.05	56 (St4)	12.11	12.20
28 (Sup5)	4.93	4.39	57 (St10)	12.23	12.35
29 (Sup10)	5.37	5.45	58 (St18)	12.73	12.73

Table 10

Correlations Between Item Placement Estimates

	Logist	MDS	Age	Lertap	Noharm
Logist					
MDS	0.99				
Age	0.99	0.99			
Lertap	-0.99	-1.00	-0.99		
Noharm	0.97	0.98	0.97	-0.98	

Logist = Logist item difficulties
MDS = Multidimensional scale values
Age = Age at which 50% of infants pass item
Lertap = Proportion of infants passing item
Noharm = Noharm item difficulties

Table 11

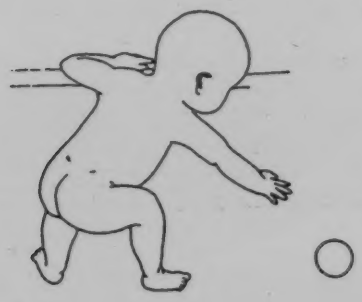
Discrimination Estimates from Various Models

	Discrimination Est.				Discrimination Est.		
Item No.	Logist	Noharm	Test graf	Item No.	Logist	Noharm	Test graf
1	5.05	****	0.28	30	5.05	46.06	2.55
2	5.05	****	2.82	31	5.05	****	2.02
3	5.05	14.70	2.53	32	5.05	30.32	2.13
4	5.05	27.66	2.48	33	5.05	30.30	2.29
5	5.05	32.12	2.52	34	5.05	38.30	2.23
6	5.05	37.10	2.39	35	5.05	42.65	2.59
7	5.05	38.97	2.17	36	5.05	42.63	2.45
8	5.05	40.63	2.52	37	5.05	43.86	2.38
9	5.05	40.86	2.17	38	5.05	48.67	2.77
10	5.05	43.33	2.46	39	5.05	45.70	2.42
11	5.05	46.82	2.63	40	5.05	46.93	2.43
12	5.05	45.28	2.46	41	5.05	48.59	2.70
13	5.05	47.58	2.61	42	5.05	43.69	2.46
14	5.05	48.43	2.72	43	5.05	****	0.40
15	5.05	48.19	2.65	44	5.05	4.10	2.36
16	5.05	48.88	2.68	45	5.05	36.81	1.97
17	5.05	47.99	2.66	46	5.05	48.08	2.84
18	5.05	44.51	2.44	47	5.05	47.65	2.84
19	5.05	44.78	2.43	48	5.05	46.21	2.73
20	5.05	46.84	2.61	49	5.05	46.54	2.78
21	5.05	40.70	2.39	50	5.05	46.48	2.77
22	5.05	****	0.00	51	5.05	44.99	2.78
23	5.05	****	1.50	52	5.05	41.99	2.78
24	5.05	8.48	2.62	53	5.05	33.94	2.78
25	5.05	26.07	2.38	54	5.05	30.10	2.79
26	5.05	35.20	2.45	55	5.05	24.65	2.63
27	5.05	34.94	2.03	56	5.05	25.41	2.71
28	5.05	38.69	2.29	57	5.05	23.67	2.43
29	5.05	42.40	2.30	58	5.05	19.49	2.53

Appendix A

Sample Items

Controlled Lowering from Standing



Weight Bearing	Weight on feet One arm support
Posture	Holds onto support with one hand
Antigravity Movement	Controlled lowering from standing

To pass this item, the infant must assume standing independently. A variety of leg postures may be observed; the legs may move symmetrically or asymmetrically. To pass this item, the movement must be controlled and the infant must not accidentally fall from standing.

PROMPT: May use toys to elicit the antigravity movements.

Sitting with Propped Arms

Weight Bearing	Weight on buttocks, legs and hands
Posture	Head up; shoulders elevated Hips flexed, externally rotated and abducted Knees flexed Lumbar and thoracic spine rounded
Antigravity Movement	Maintains head in midline Supports weight on arms briefly

PROMPT: Examiner places the infant in sitting.

Hands to Knees

Weight Bearing	Weight symmetrically distributed on head, trunk and pelvis
Posture	Hips abducted, externally rotated Knees flexed Pelvis neutral moving towards a posterior tilt
Antigravity Movement	Turns head easily side to side Chin tuck Reaches hand(s) to knees Abdominals active May fall to side by lifting legs

It is important to observe active abdominals. If the legs are widely abducted and resting on the abdomen passively, the infant would not pass this item. Hypotonic infants often display this passive position.

Extended Arm Support

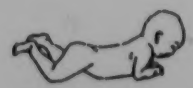
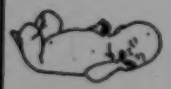
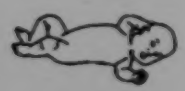
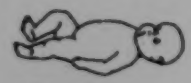
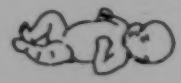
Weight Bearing	Weight on hands, lower abdomen and thighs
Posture	Arms extended Elbows in front of shoulders Legs approaching neutral position
Antigravity Movement	Chin tuck and chest elevated Flexion and extension of knees; may play with feet together Lateral weight shift

The infant may also push backwards in this position.

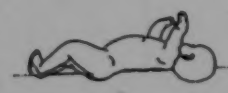
Appendix B

Score Sheets

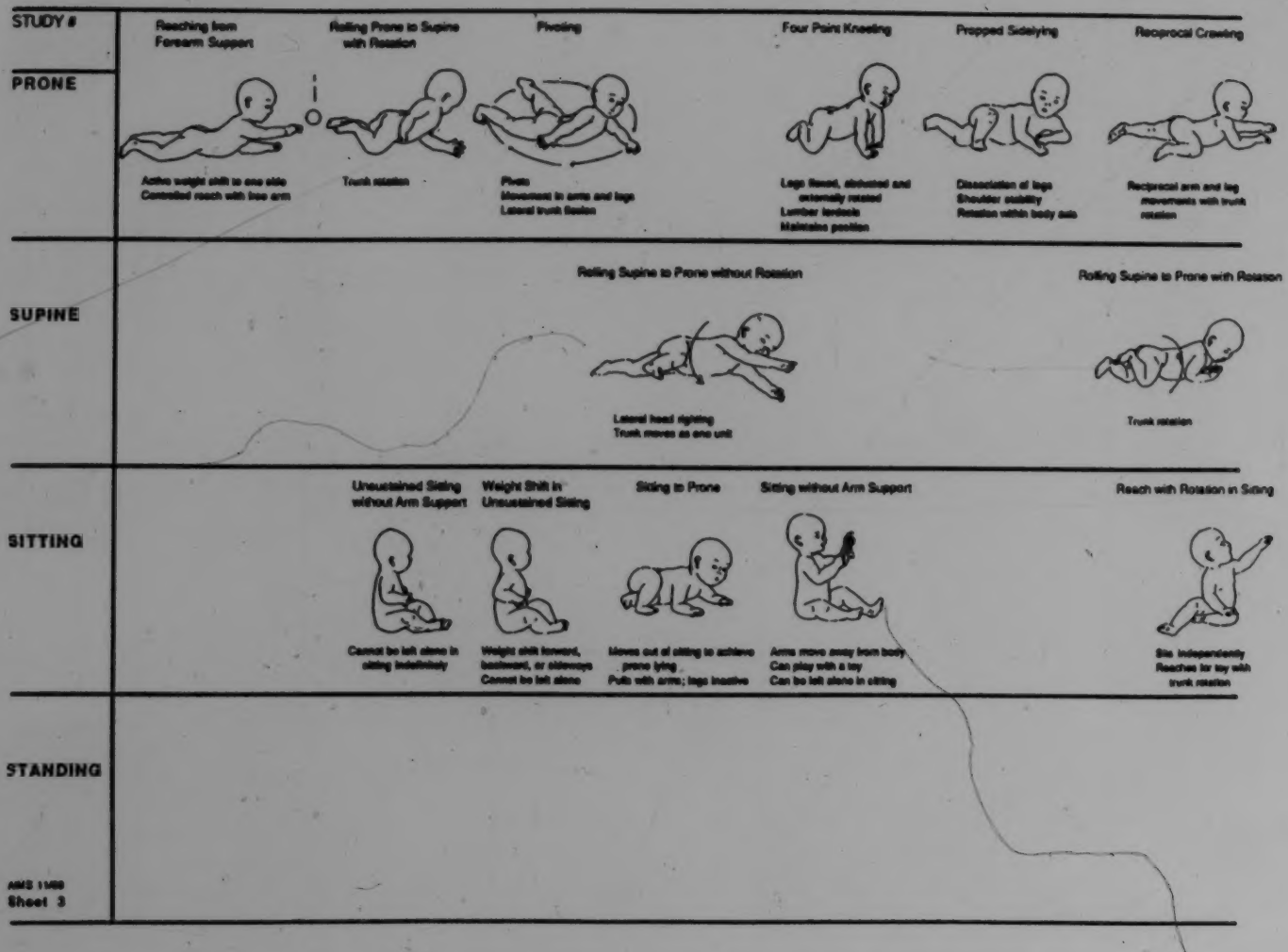
SCORE SHEETS: AIMS RELIABILITY AND VALIDITY STUDY

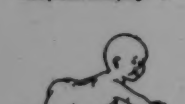
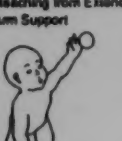
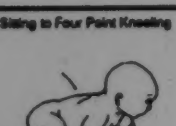
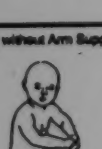
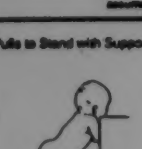
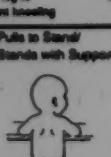
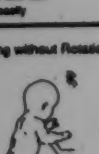
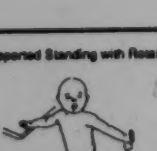
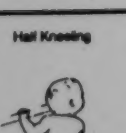
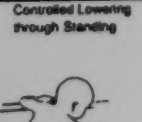
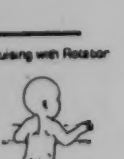
STUDY #					
PRONE	Prone Lying Physiological flexion Turns head to clear nose from surface	Prone Lying Lifts head asymmetrically to 45° Cannot maintain head in midline	Prone Prop  Elbows behind shoulders Unsustained head raising to 45°	Forearm Support Lifts and maintains head past 45° Elbows in line with shoulders Chest elevated	Prone Mobility Head up to 90° Uncontrolled weight shift
SUPINE	Supine Lying  Physiological flexion Head rotation; mouth to hand Random arm and leg movements	Supine Lying  Head rotation towards midline Non-obligatory ATNR	Supine Lying  Head in midline Moves arms but unable to bring hands to midline	Supine Lying  Neck flexors active - chin tuck Brings hands to midline	
SITTING	Sitting with Support Lifts and maintains head in midline briefly				Pull to Sit Chin tuck; head in line or in front of body
STANDING	Supported Standing May have intermittent hip and knee flexion	Supported Standing Head in line with body Hips behind shoulders Variable movement of legs			

AIMS 11/89
Sheet 1

STUDY #				
PRONE	Forearm Support Elbows in front of shoulders Active chin tuck with neck elongation	Rolling Prone to Supine Movement initiated by head Trunk moves as one unit	Extended Arm Support Arms extended Chin tuck and chest elevation Lateral weight shift	Swimming Active extension pattern
SUPINE	Hands to Knees Chin tuck Reaches hands to knees Abdominals active		Active Extension  Pushes into extension with legs	Hands to Feet Can maintain legs in mid-range Pelvis mobility present
SITTING	Sitting with Propped Arms Maintains head in midline Supports weight on arms briefly		Unsupported Sitting Scapular adduction and humeral extension Topples forward or to side	Sitting with Arm Support Thoracic spine extended Head movements free from trunk; propped on extended arms
STANDING	Supported Standing Hips in line with shoulders Active control of trunk Variable movements of legs			

AIMS 11/90
Sheet 2


STUDY #
PRONE
Reaching from Forearm Support
Active weight shift to one side
Controlled reach with free arm
Rolling Prone to Supine with Rotation
Trunk rotation
Pivoting
Pivots
Movement in arms and legs
Lateral trunk flexion
Four Point Kneeling
Legs flexed, abducted and externally rotated
Lumbar lordosis
Maintains position
Propped Sidelying
Dissociation of legs
Shoulder stability
Rotation within body axis
Reciprocal Crawling
Reciprocal arm and leg movements with trunk rotation
SUPINE
Rolling Supine to Prone without Rotation
Lateral head righting
Trunk moves as one unit
Rolling Supine to Prone with Rotation
Trunk rotation
SITTING
Unsustained Sitting without Arm Support
Cannot be left alone in sitting indefinitely
Weight Shift in Unsustained Sitting
Weight shift forward, backward, or sideways
Cannot be left alone
Sitting to Prone
Moves out of sitting to achieve prone lying
Pulls with arms; legs inactive
Sitting without Arm Support
Arms move away from body
Can play with a toy
Can be left alone in sitting
Reach with Rotation in Sitting
Sits independently
Reaches for toy with trunk rotation
STANDING
AMS 1148
Sheet 3

STUDY #						
PRONE	Four Point Kneeling to Sitting or Half Sitting	Reciprocal Creeping	Four Point Kneeling	Modified Four Point Kneeling	Reaching from Extended Arm Support	Reciprocal Creeping with Rotation
						
	Plays in and out of position May get to sitting	Legs abducted and externally rotated Lumbar lordosis; weight shift side to side with lateral trunk flexion	Hips aligned under pelvis; flattening of lumbar spine	Plays in position May move forward	Reaches with extended arm Trunk rotation	Lumbar spine flat Moves with trunk rotation
SUPINE						
SITTING	Sitting to Four Point Kneeling	Sitting without Arm Support				
						
	Actively lifts pelvis, buttocks, and unweighted leg to assume four point kneeling	Position of legs varies Infant moves in and out of positions easily				
STANDING	Pulls to Stand with Support	Pulls to Stand/ Stands with Support	Cruising without Rotation	Supported Standing with Rotation	Half Kneeling	Controlled Lowering through Standing	Cruising with Rotation
							
	Pushes down with arms and extends knees	Pulls to stand; shifts weight from side to side	Cruises sideways without rotation	Rotation at trunk and pelvis	May assume standing or play in position	Controlled lowering from standing	Cruises with rotation

Abt5 1149
Sheet 1

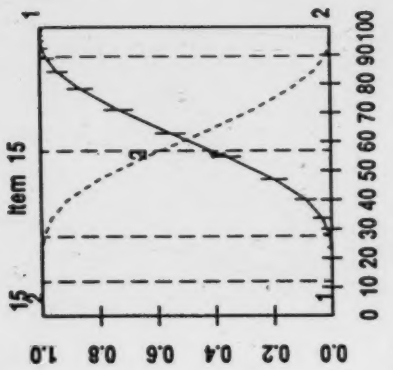
STUDY #								
PRONE								
SUPINE								
SITTING								
STANDING	Stands Alone Stands alone momentarily Balance reactions in feet	Early Stepping Walks independently; moves quickly with short steps; lateral flexion of trunk	Standing from Quadruped Position Pushes quickly with knees to get to standing	Standing from Modified Squat Moves from squat to standing with controlled flexion and extension of hips and knees	Walks Alone Walks independently	Squat Maintains position by balance reactions in feet and in position of trunk		
AMS 1100 Sheet 5								

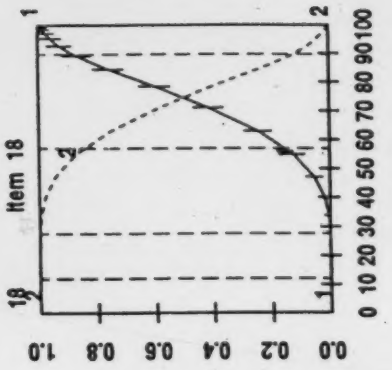
Appendix C

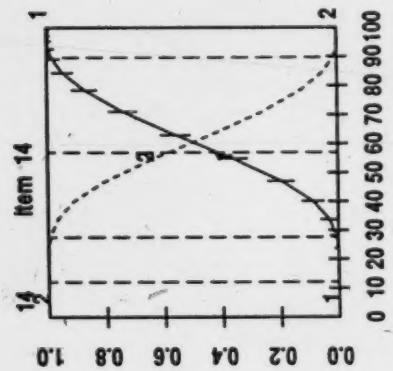
Item Characteristic Curves

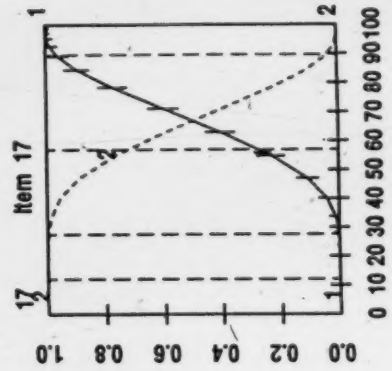
Item 1
Item 2
Item 3
Item 4
Item 5
Item 6
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90 100

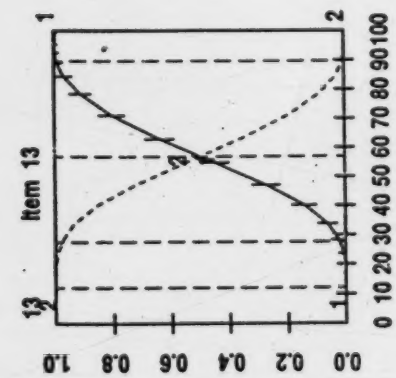
Item 7
Item 8
Item 9
Item 10
Item 11
Item 12
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

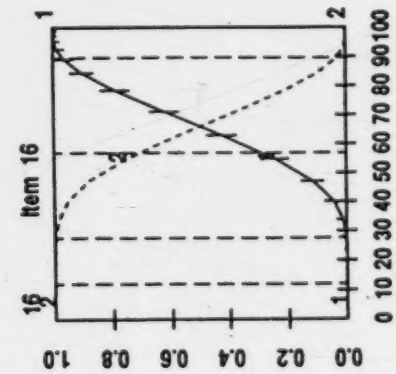

Item 15
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

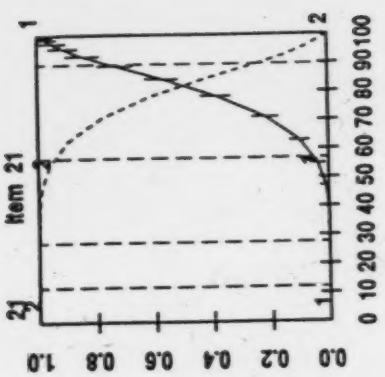

Item 18
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

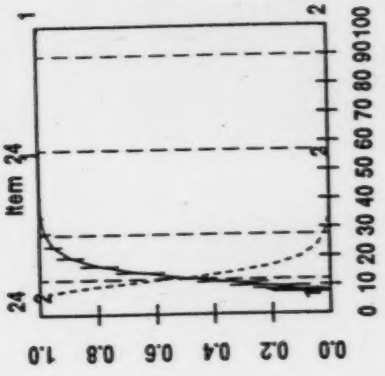

Item 14
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

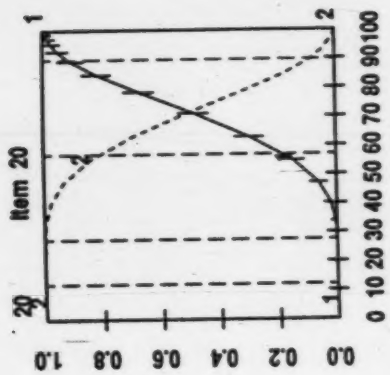

Item 17
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

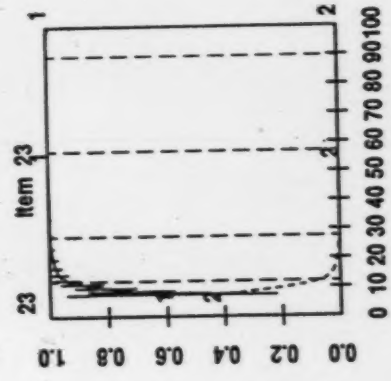

Item 13
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

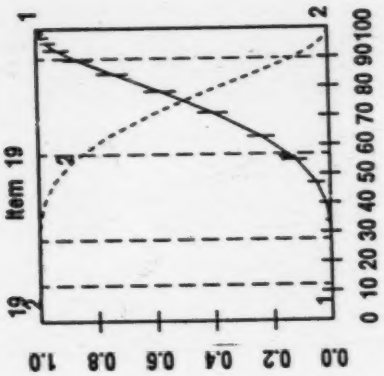

Item 16
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

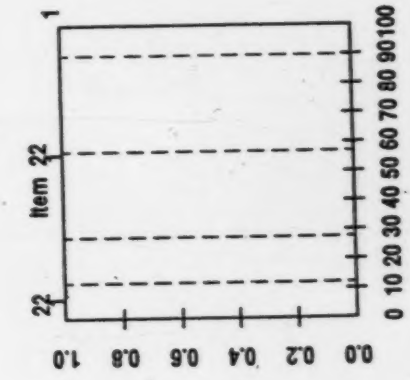

Item 21
21
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

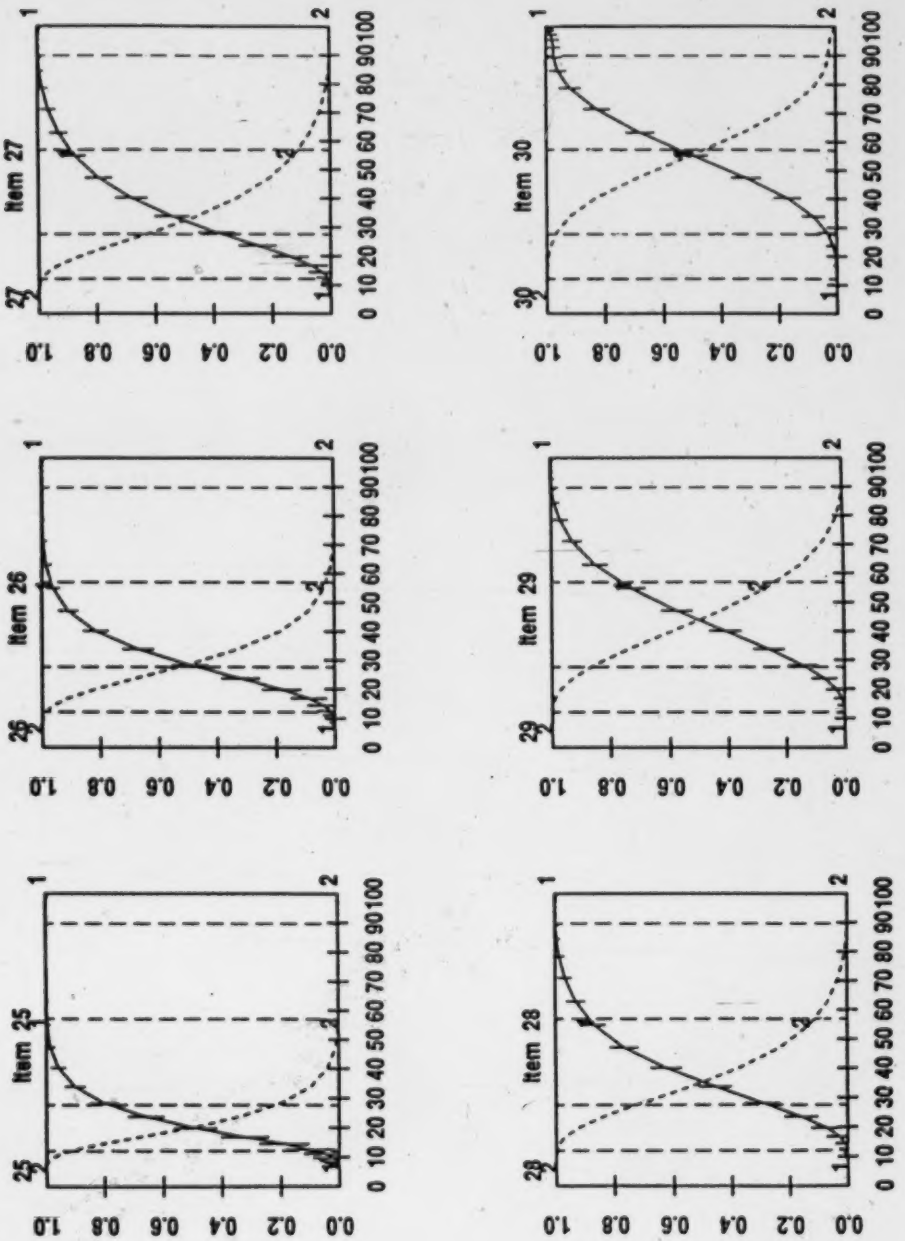

Item 24
24
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

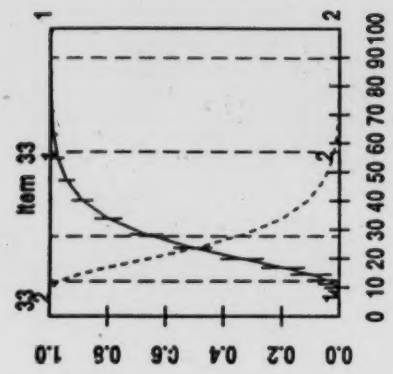

Item 20
20
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

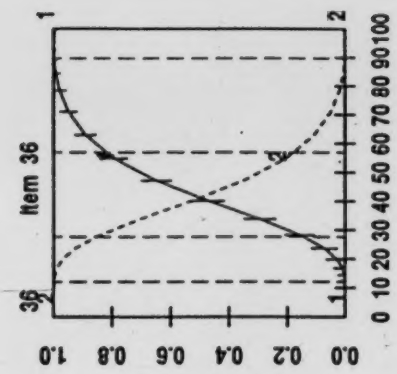

Item 23
23
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

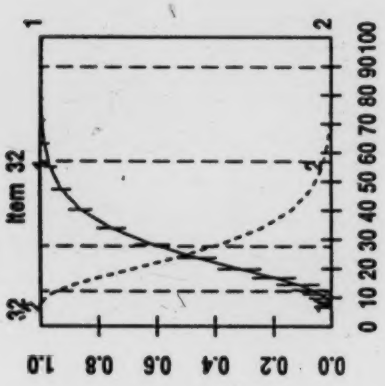

Item 19
19
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

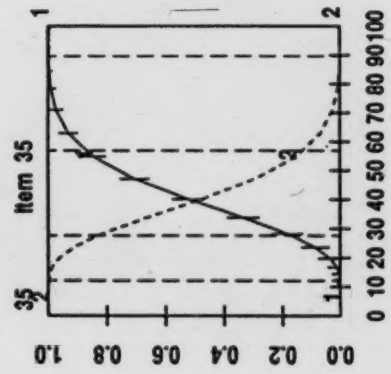

Item 22
22
1
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

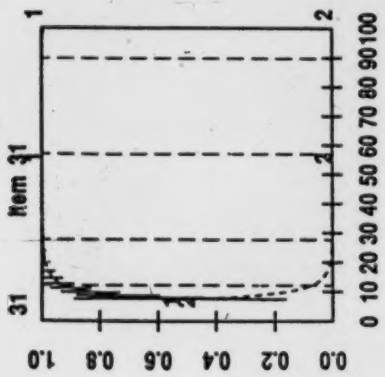

Item 25
Item 26
Item 27
Item 28
Item 29
Item 30
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90 100

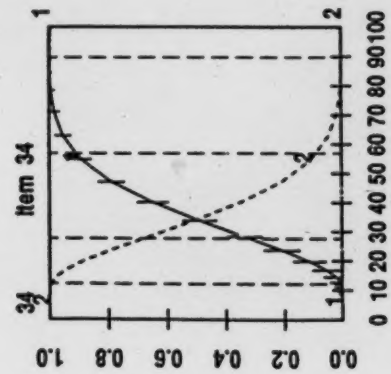

Item 33
33
1
2
0 10 20 30 40 50 60 70 80 90100
0.0 0.2 0.4 0.6 0.8 1.0

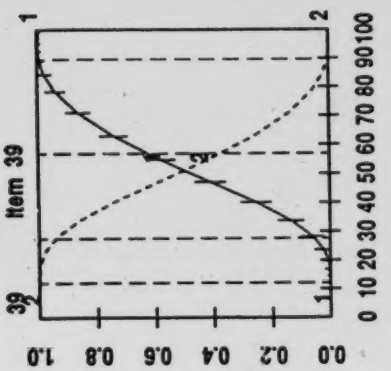

Item 36
36
1
2
0 10 20 30 40 50 60 70 80 90100
0.0 0.2 0.4 0.6 0.8 1.0

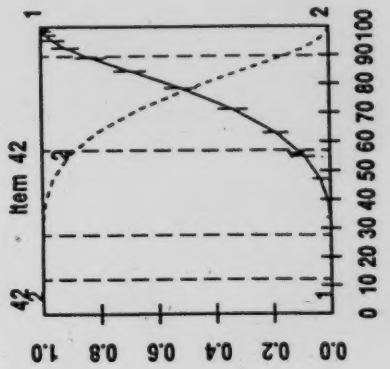

Item 32
32
1
2
0 10 20 30 40 50 60 70 80 90100
0.0 0.2 0.4 0.6 0.8 1.0

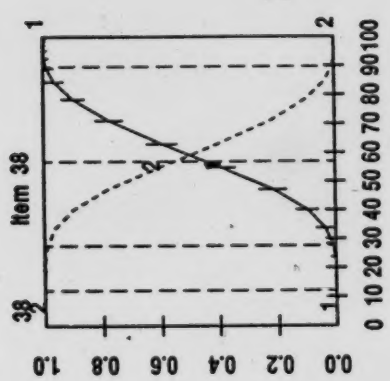

Item 35
35
1
2
0 10 20 30 40 50 60 70 80 90100
0.0 0.2 0.4 0.6 0.8 1.0

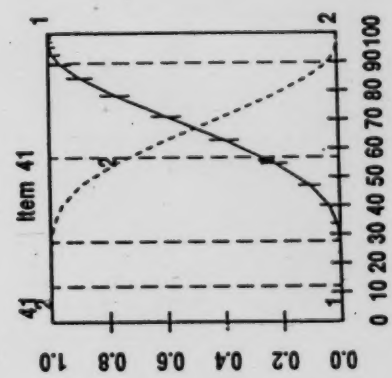

Item 31
31
1
2
0 10 20 30 40 50 60 70 80 90100
0.0 0.2 0.4 0.6 0.8 1.0

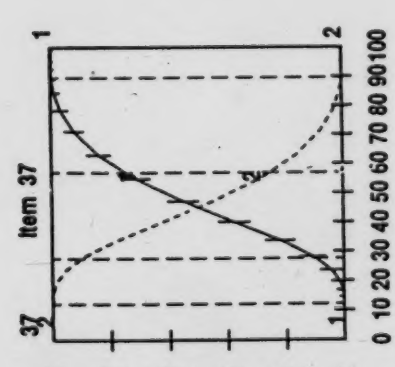

Item 34
34
1
2
0 10 20 30 40 50 60 70 80 90100
0.0 0.2 0.4 0.6 0.8 1.0

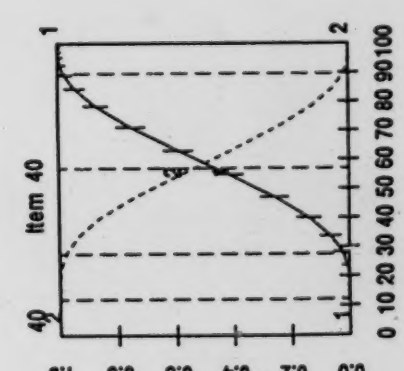

Item 39
39
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

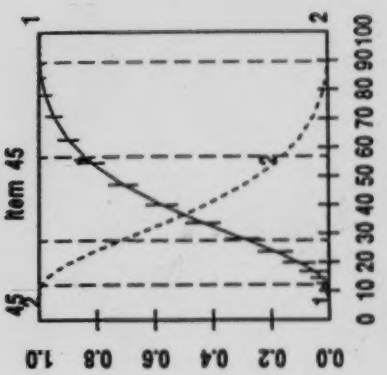

Item 42
42
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

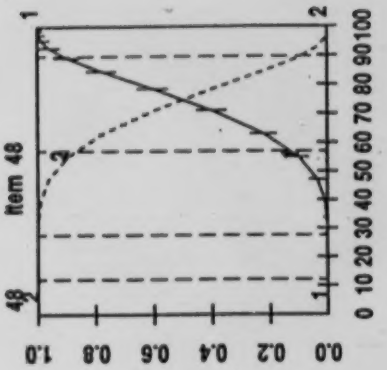

Item 38
38
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

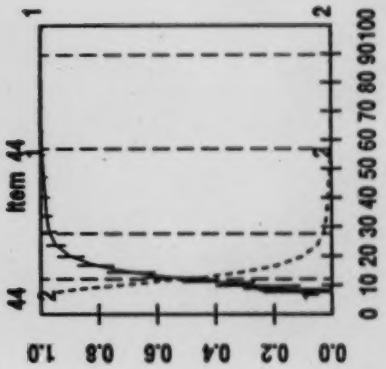

Item 41
41
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

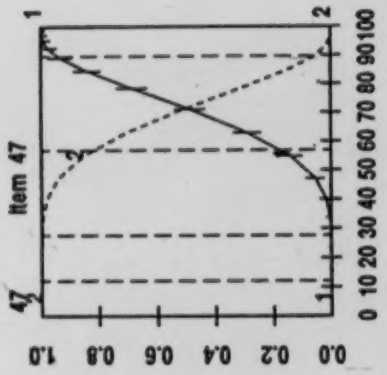

Item 37
37
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

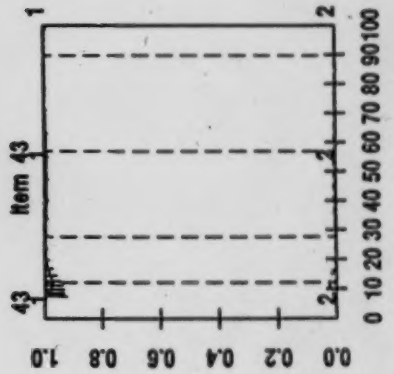

Item 40
40
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

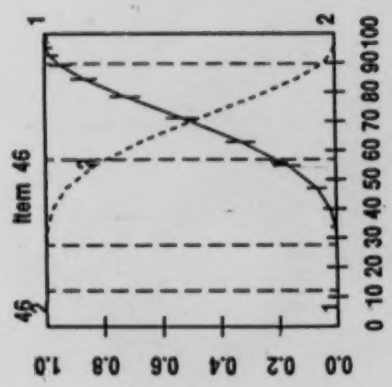

Item 45
45
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100


Item 48
48
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

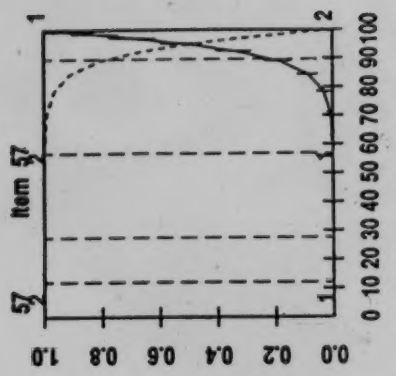

Item 44
44
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

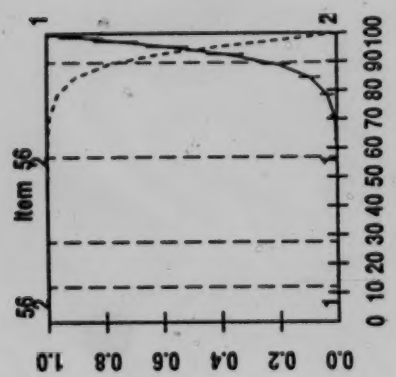

Item 47
47
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

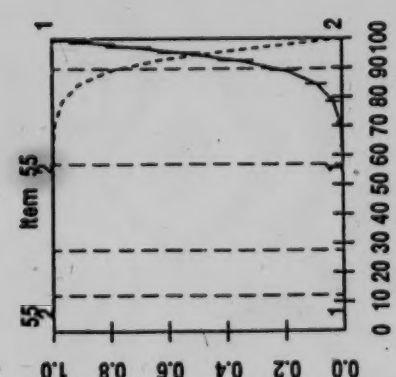

Item 43
43
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

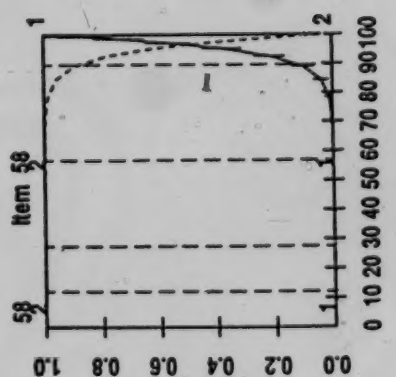

Item 46
46
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90100

Item 51
Item 54
Item 50
Item 53
Item 49
Item 52
0 10 20 30 40 50 60 70 80 90100
0.0 0.2 0.4 0.6 0.8 1.0


Item 57
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90 100


Item 56
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90 100


Item 55
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90 100

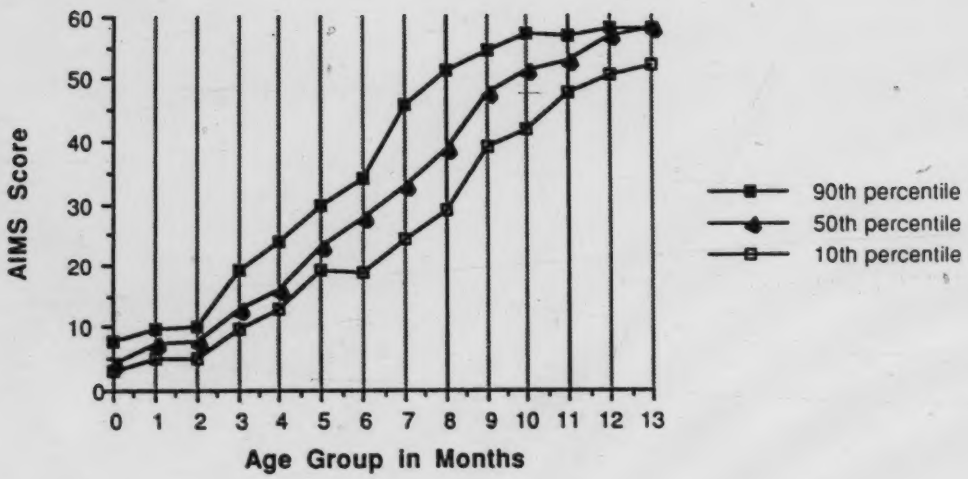

Item 58
1
2
0.0 0.2 0.4 0.6 0.8 1.0
0 10 20 30 40 50 60 70 80 90 100

Appendix D

Sample Norms Graph

Appendix D

Sample Norms Graph



Appendix E

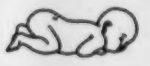
Revised Score Sheets

Alberta Infant Motor Scale

STUDY #

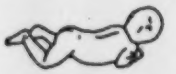
PRONE

Prone Lying



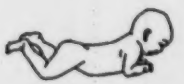
Physiological flexion
Turns head to clear nose from surface

Prone Lying



Lifts head asymmetrically to 45°
Cannot maintain head in midline

Prone Prop



Elbows behind shoulders
Unsustained head raising to 45°

Forearm Support

Lifts and maintains head past 45°
Elbows in line with shoulders
Chest elevated

Prone Mobility

Head up to 90°
Unsupported weight shift

SUPINE

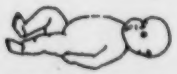
Supine Lying

Physiological flexion
Head rotation/ attends to hand
Random arm and leg movements

Supine Lying

Head rotation towards midline
Non-obligatory ATNR

Supine Lying



Head in midline
Moves arms but unable to bring hands to midline

Supine Lying

Neck flexors active - chin tuck
Brings hands to midline

SITTING

Sitting with Support

Lifts and maintains head in midline briefly

Sitting with Propped Arms

Maintains head in midline
Supports weight on arms briefly

Pull to Sit

Chin tucks head in line or in front of body

STANDING

Supported Standing

May have intermittent hip and knee flexion

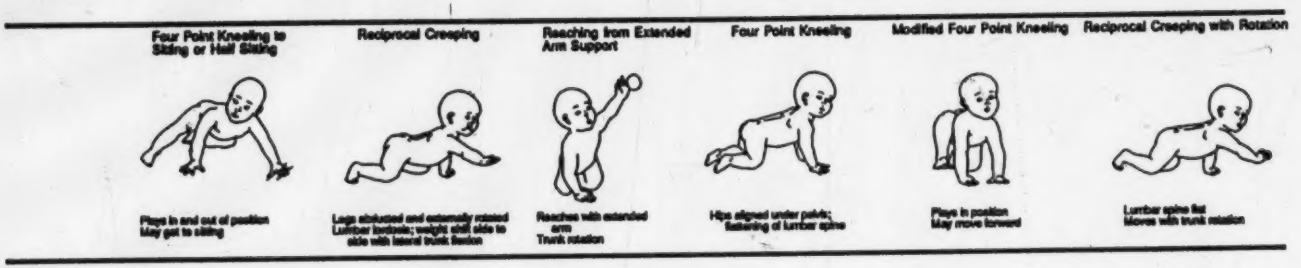
Supported Standing

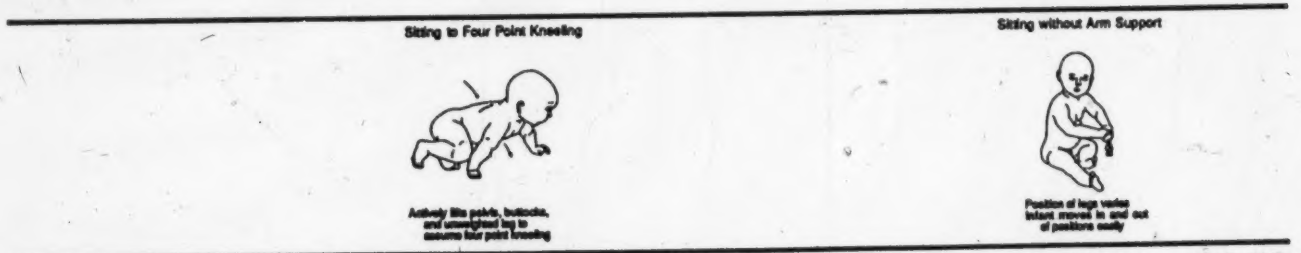
Head in line with body
Hips behind shoulders
Variable movement of legs

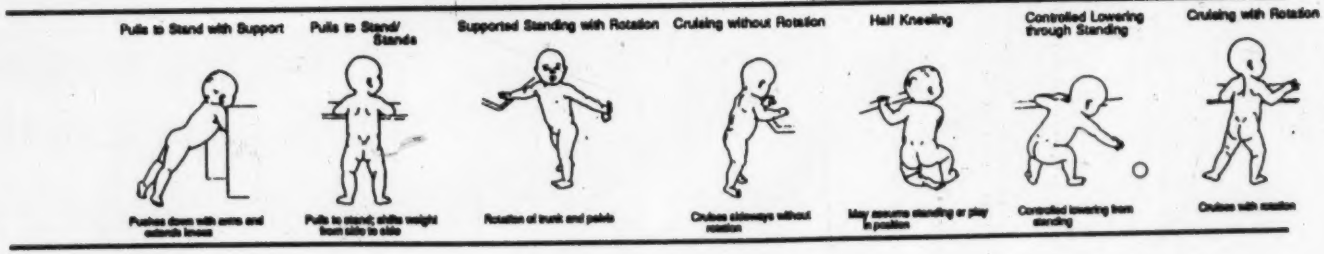
AIMS 07/91
Sheet 1

Forearm Support
Extended Arm Support
Rolling Prone to Supine
Swimming
Elbows in front of shoulders
Active chin lift with neck elongation
Arms extended
Chin held and chest elevation
Lateral weight shift
Movement initiated by head
Trunk moves as one unit
Active swimmer pattern
Hands to Knees
Active Extension
Hands to Feet
Chin held
Extended hands to knees
Abdominals active
Pushes into extension with legs
Can extricate legs in mid-range
Pulls to midline passed
Unsupported Sitting
Sitting with Arm Support
Unsupported Sitting without Arm Support
Supported Standing

Reaching from Forearm Support
Pivoting
Rolling Prone to Supine with Rotation
Four Point Kneeling
Propped Sidelying
Reciprocal Crawling
Rolling Supine to Prone without Rotation
Rolling Supine to Prone with Rotation
Weight Shift in Unsupported Sitting
Sitting without Arm Support
Reach with Rotation in Sitting
Sitting to Prone


Four Point Kneeling to Sitting or Half Sitting
Plays in and out of position
May get to sitting
Reciprocal Creeping
Legs abducted and externally rotated
Lumbar lordosis; weight shift side to side with lateral trunk flexion
Reaching from Extended Arm Support
Reaches with extended arm
Trunk rotation
Four Point Kneeling
Hips aligned under pelvis; flattening of lumbar spine
Modified Four Point Kneeling
Plays in position
May move forward
Reciprocal Creeping with Rotation
Lumbar spine flat
Moves with trunk rotation

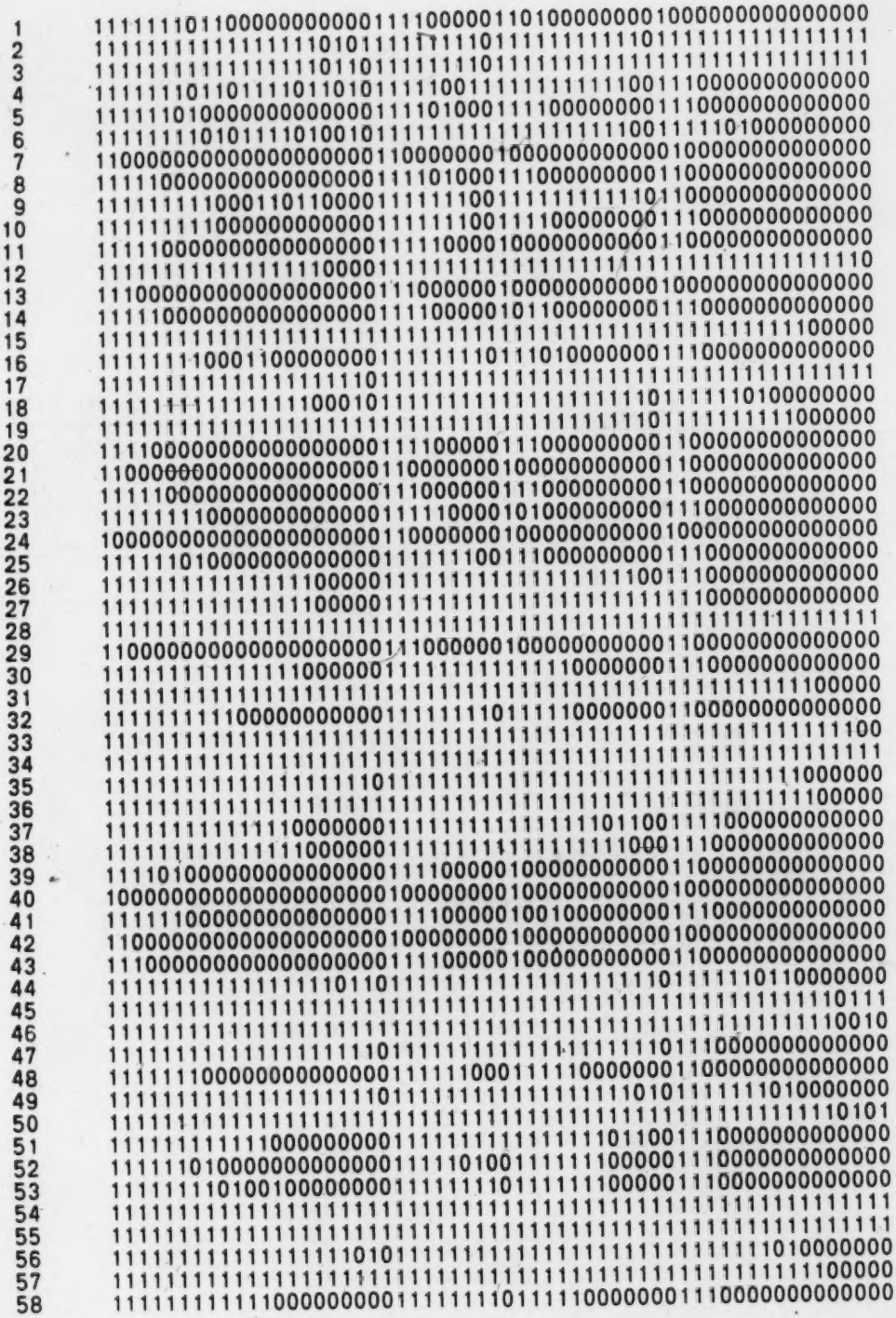

Sitting to Four Point Kneeling
Actively lifts pelvis, buttocks, and unweighted leg to assume four point kneeling
Sitting without Arm Support
Position of legs varies
Infant moves in and out of position easily

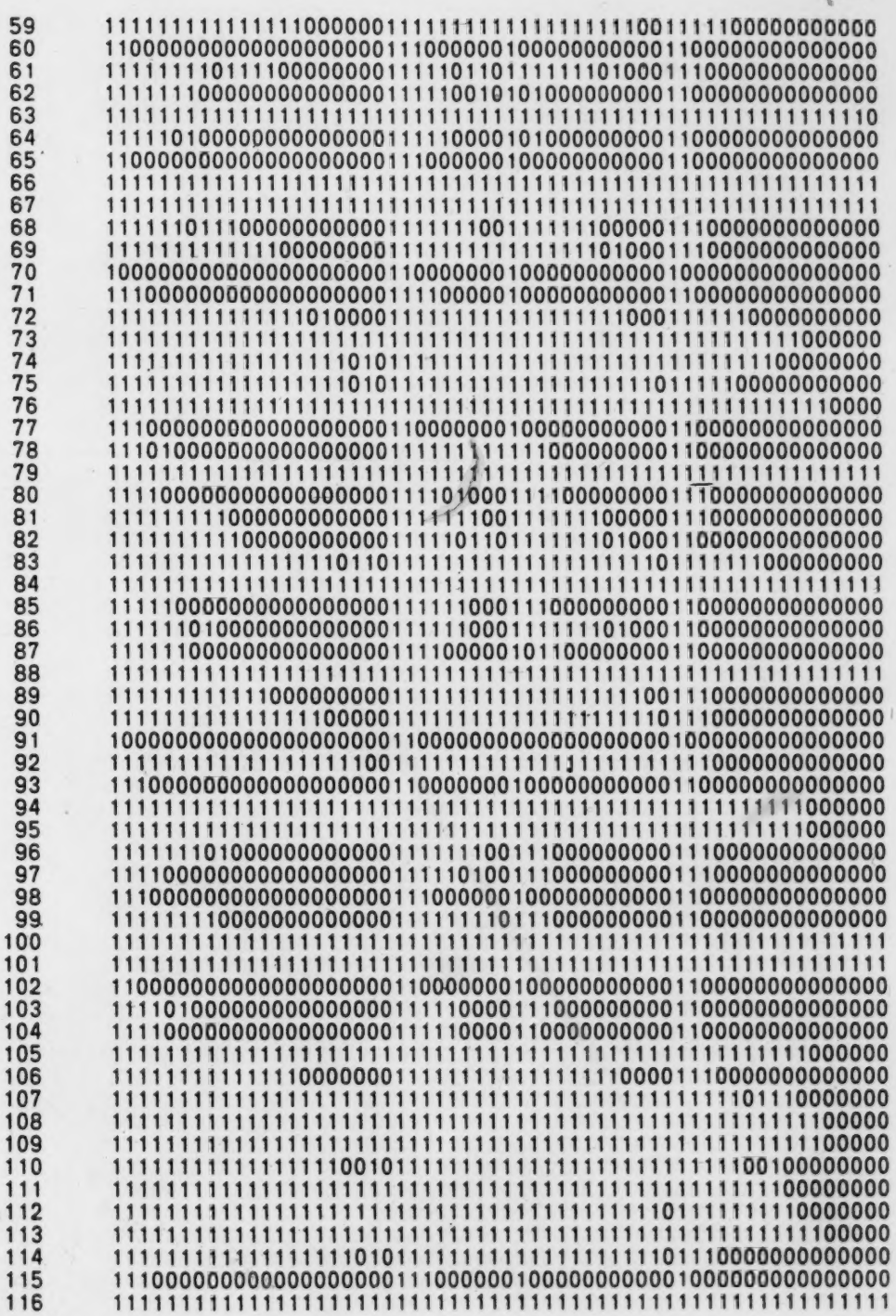

Pulls to Stand with Support
Pushes down with arms and extends knees
Pulls to Stand/ Stands
Pulls to stand; shifts weight from side to side
Supported Standing with Rotation
Rotation of trunk and pelvis
Cruising without Rotation
Cruises sideways without rotation
Half Kneeling
May assume standing or play in position
Controlled Lowering through Standing
Controlled lowering from standing
Cruising with Rotation
Cruises with rotation

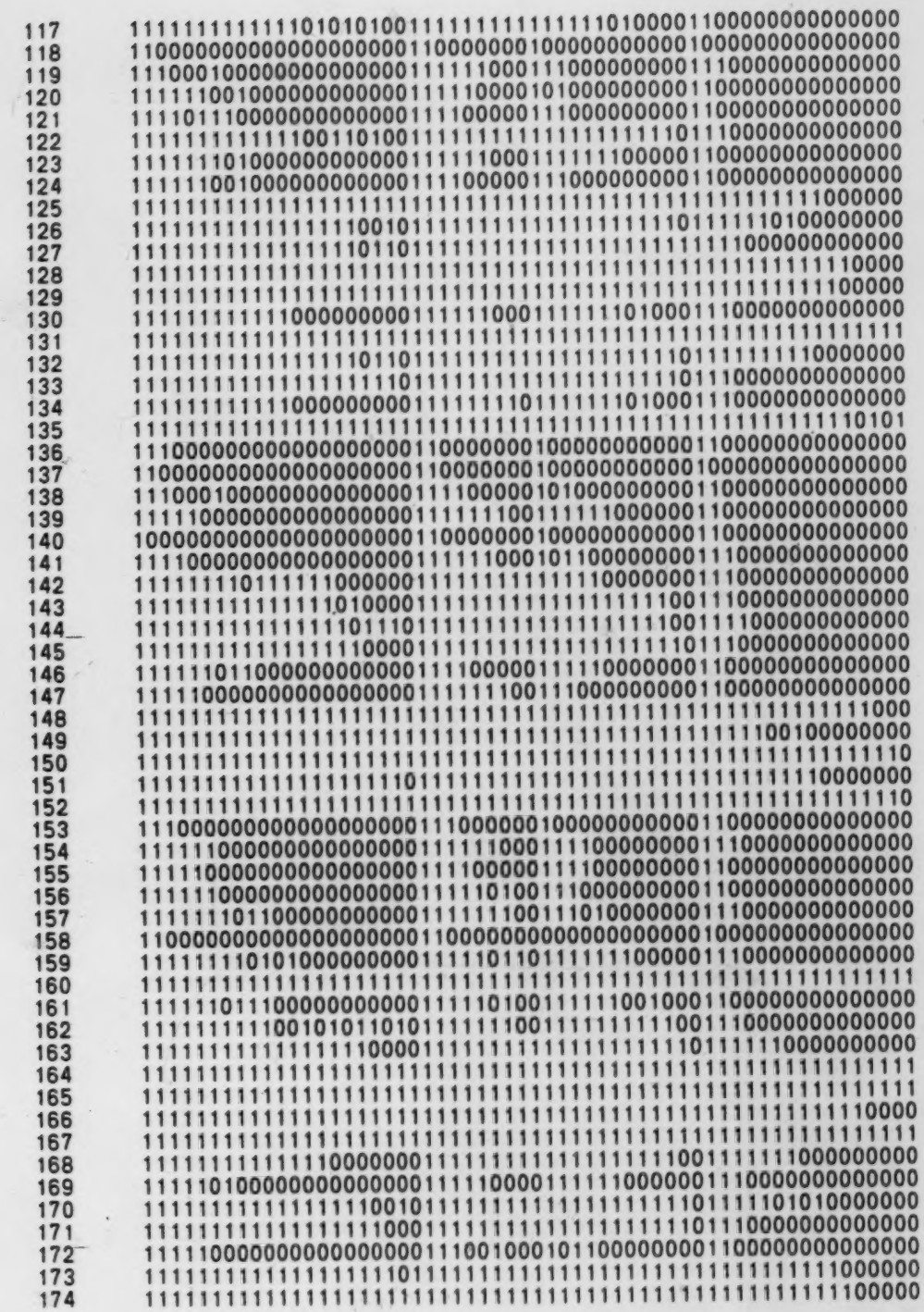
Stands Alone
Early Stepping
Standing from Modified Squat
Standing from Quadruped Position
Walks Alone
Squat

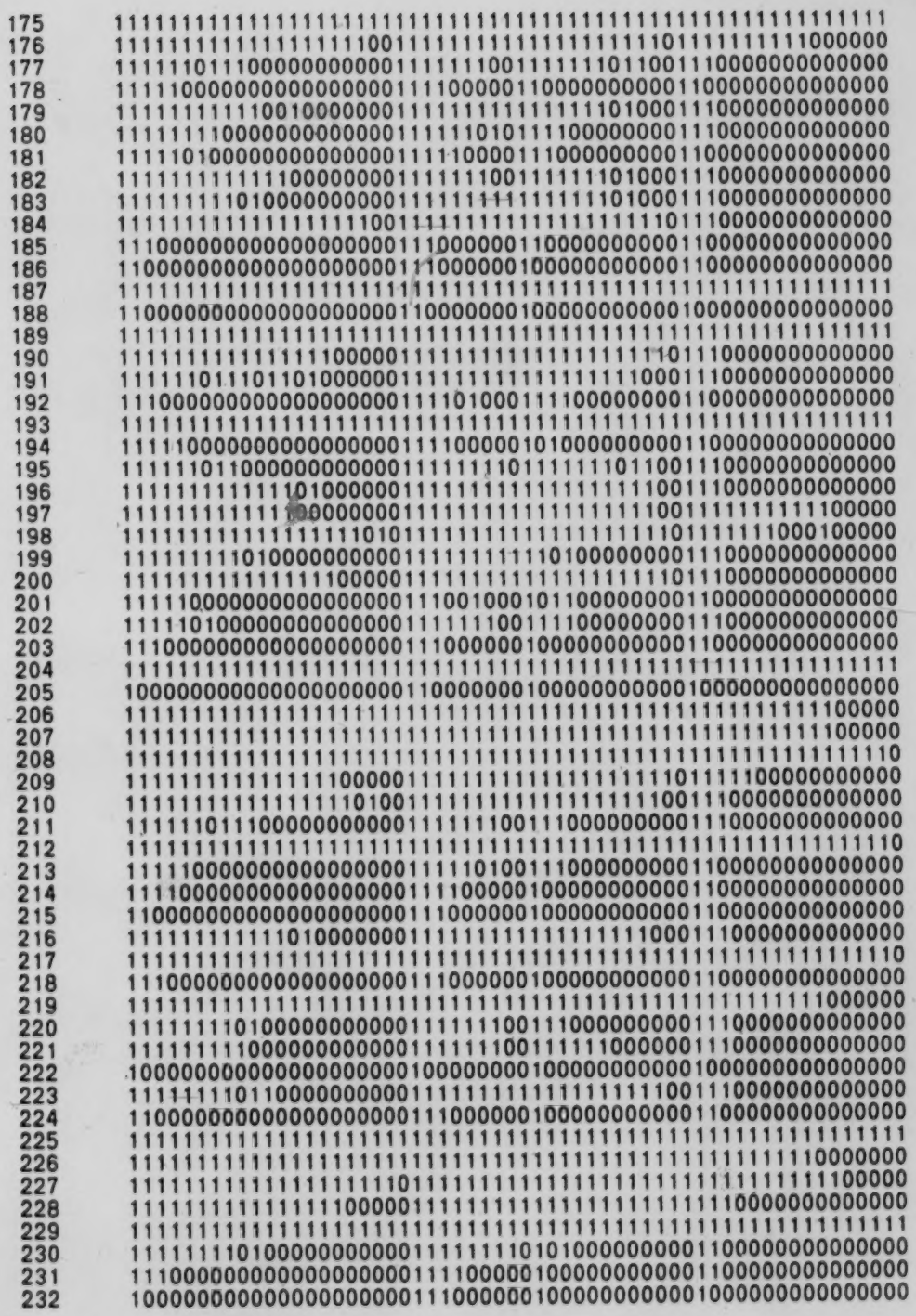
Appendix F

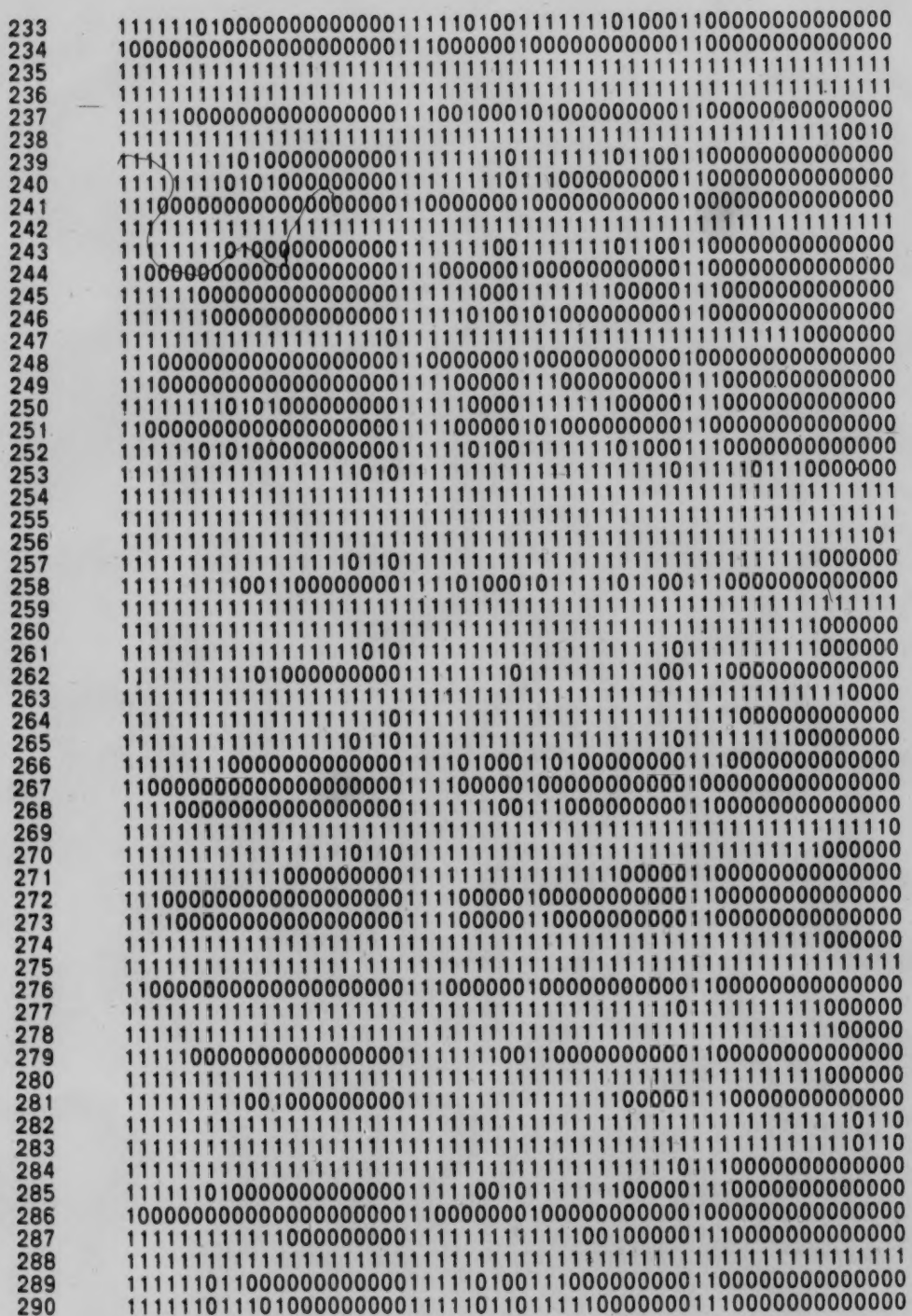
0/1 Data Set for 479 Infants

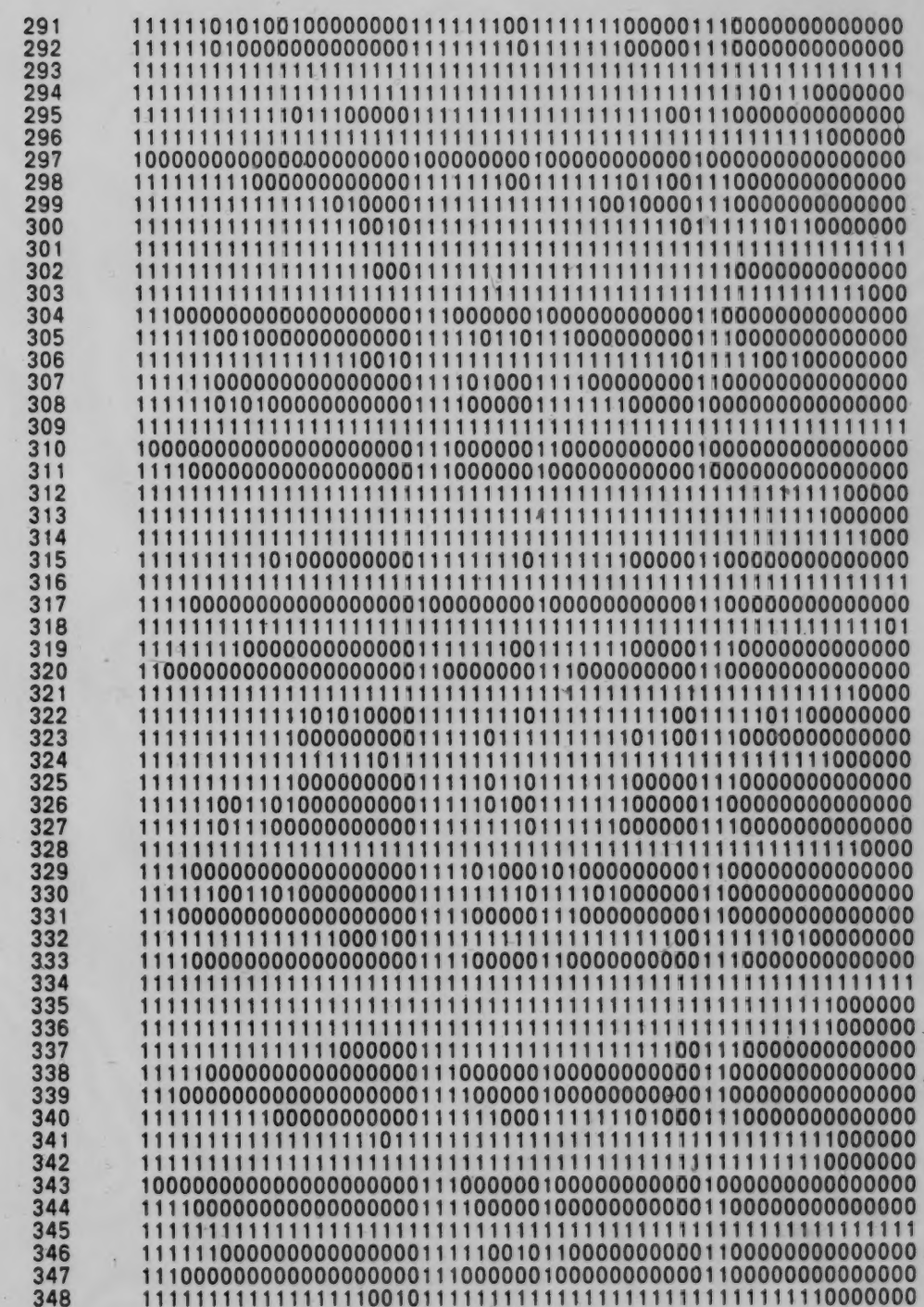


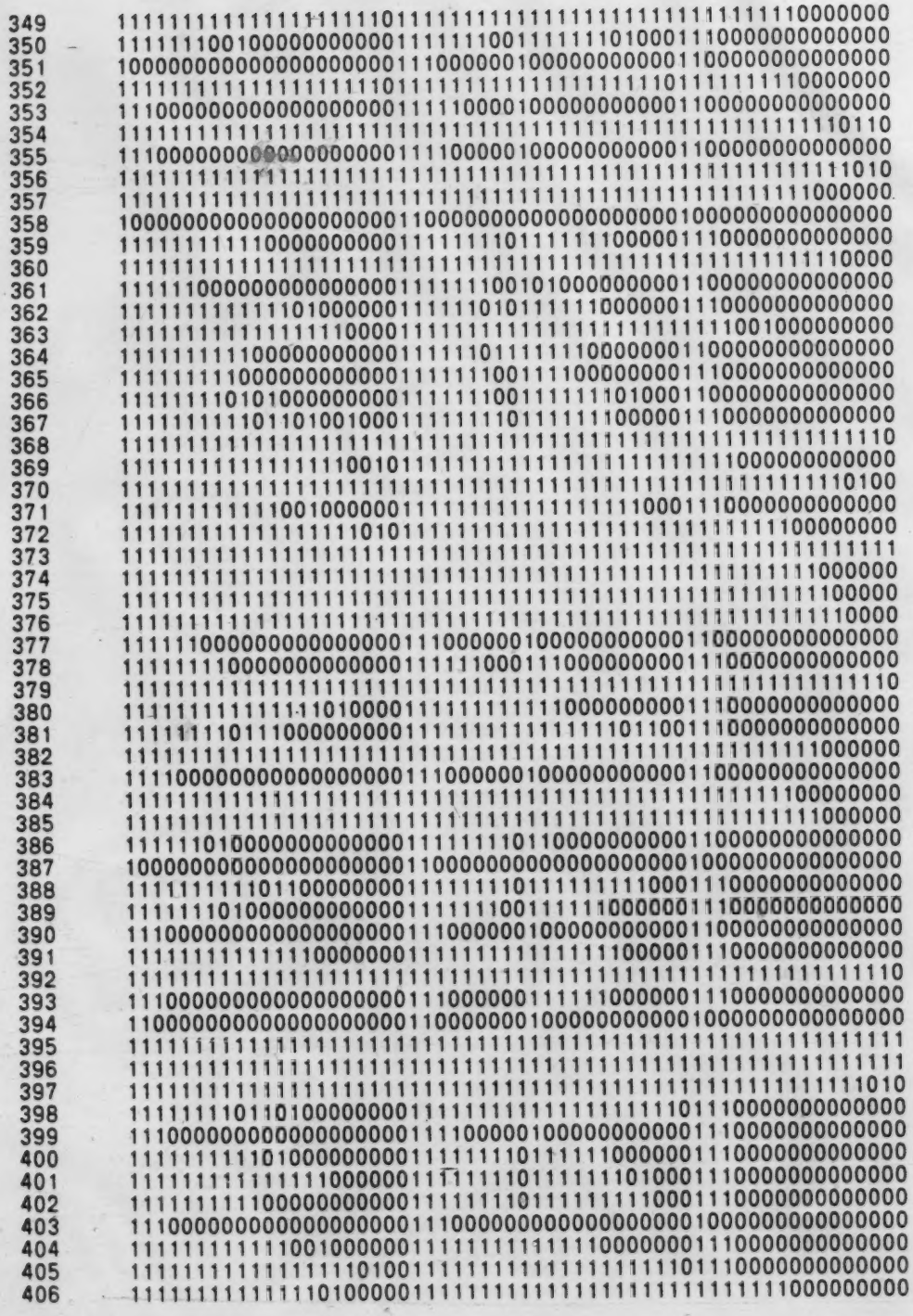


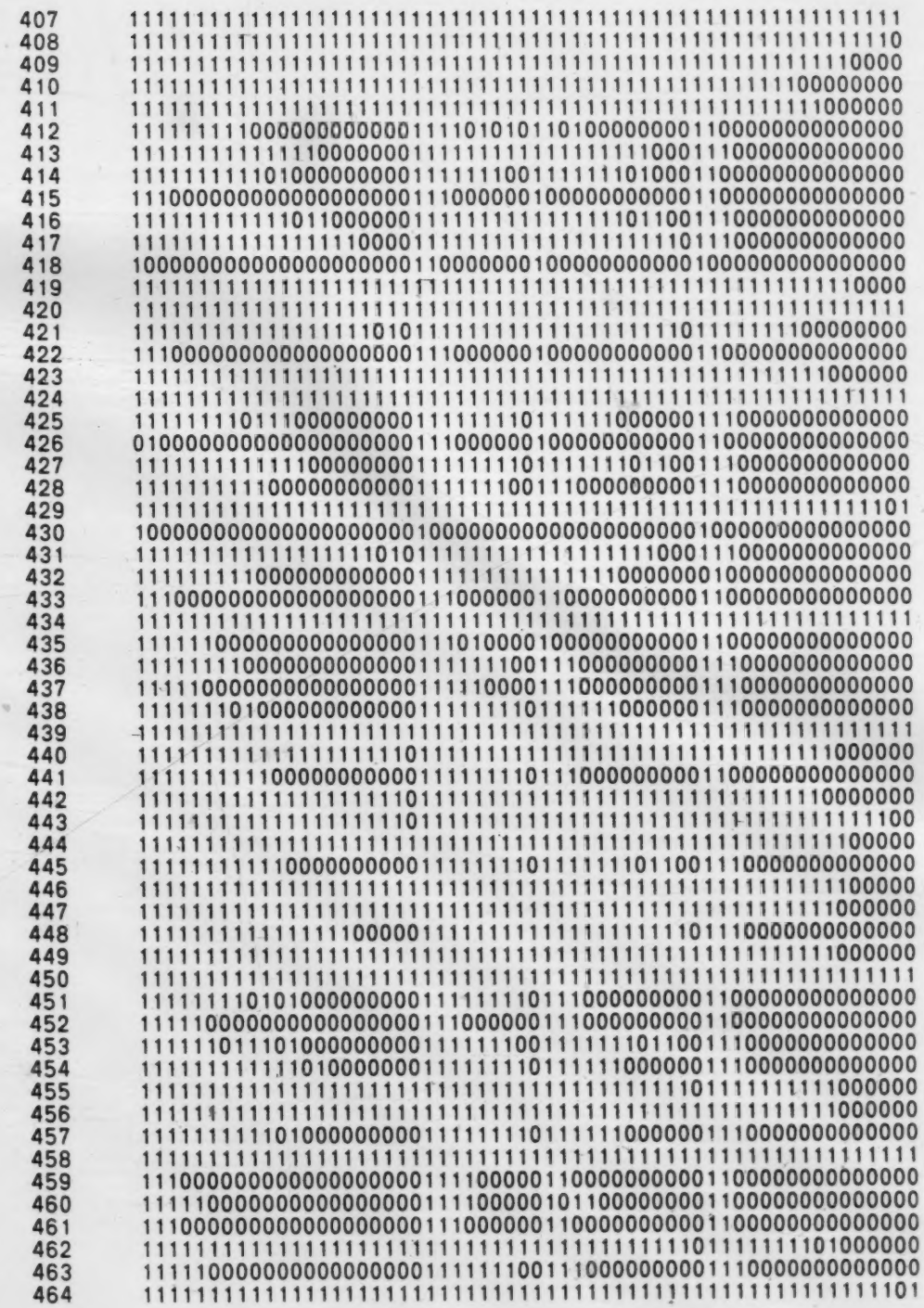












```
465    11111111111100000000011111111011111110100011100000000000000
466    11111000000000000000011111000010100000000011000000000000000
467    11000000000000000000011000000010000000000011000000000000000
468    11111000000000000000011110000010100000000011000000000000000
469    11111111010100000000011111100011111110110011100000000000000
470    11110000000000000000011100000010000000000011000000000000000
471    11111111111100000000011111111111111110000011100000000000000
472    11111111111111111111101111111111111111111111111111111110000
473    11111000000000000000011110100011100000000011000000000000000
474    11111111111111111011011111111111111111111011111010100000000
475    11111101010000000000011111110011111111100110000000000000000
476    11111111111100000000011111111111111110100011100000000000000
477    11111111111111111111011111111111111111111111111111111000000
478    11111111110110000000011111001110110100000011000000000000000
479    11111111100000000000011111000011000000000011000000000000000
```